

مدخل حديث للإحصاء والاحتمالات

إعداد
دكتورة / ثروت محمد عبد المنعم




مكتبة الأنجلو المصرية

مدخل حديث للإحصاء والاحتمالات

إعداد

د. ثروت محمد عبد المنعم

أستاذ مساعد بكلية العلوم بالدمام

قسم الرياضيات

المملكة العربية السعودية

الطبعة الأولى ١٤٢١ هـ - ٢٠٠٠ م



مكتبة الأنجلو المصرية
١٦٥ شارع محمد فريد - القاهرة

بسم الله الرحمن الرحيم

وما أوتيتم من العلم إلا قليلا

(صدق الله العظيم)

الإهداء

إلى أمي التي علمتني العطاء.

إلى والذي الذي علمني الخلق الكريم.

إلى أستاذي أ.د. أحمد حسن المولايي الذي علمني أن الثقة بالنفس أول خطوات النجاح.

إلى أستاذي أ.د. سمير كامل عاشور و أستاذتي أ.د. الهام شكري اللذان علماني أن تشجيع الأستاذ

لتلميذه دافع قوي له على التقدم.

إلى أستاذي المرحوم أ.د. علي يحيى الذي علمني أنه لا يوجد صعب في العلم. إلى أختي فهم سندی في

الحياة

إلى أخواي في الله د. سارة و د. أميرة اللتان علمتاني أن الأخوة ليست فقط في الرحم.

إلى كل من شجعني في حياتي و أعطاني دفعة نحو الأمام.

بسم الله الرحمن الرحيم

تقدم

الحمد لله رب العالمين والصلاة والسلام على أشرف المرسلين محمد وعلى صحبه أجمعين وبعد...
يعون الله سبحانه وتعالى تقدم هذا الكتاب للمكتبة العربية وللباحثين العرب في المجالات المختلفة تلبية لنداء التعريب الذي يتناهى الكثير من العلماء و المثقفين . يهدف هذا الكتاب إلى تقديم الطرق الإحصائية للدارسين والطلاب المتخصصين في الإحصاء مع عتد الحوض في الكثير من الإشتاقات الرياضية وكذلك البعد عن التعقيد في النظريات الإحصائية والتركيز على الفهم والتطبيق . وقد استندت في الشرح النظري على الكثير من الأمثلة التطبيقية المدعمة بالتمارين التي تلي كل فصل في الكتاب لتؤكد الفهم واختيار القدرة على التطبيق .

معظم الكتب الإحصائية تبدأ بعرض الطرق الوصفية ثم بموضوع الاحتمالات كجزء منفصل ثم بطرق الاستدلال الإحصائي . يقدم هذا الكتاب المحتويات بأسلوب مختلف حيث يبدأ بموضوع الاحتمالات ثم يربط مباشرة بالطرق الإحصائية سواء في مجال الإحصاء الوصفي أو الإستدلال . هذا وقد راعت أن تكون المعادلات بالرموز اللاتينية وذلك للتيسر على القارئ في الاستفادة من المراجع الأجنبية . كما استعت في إعداد هذا الكتاب أولاً بخبرتي الطويلة في تدريس الإحصاء وفي الأبحاث والاستشارات الإحصائية وثانياً بكثير من المراجع العربية والأجنبية والمطعة أسمائهم في نهاية الكتاب . هذا وقد تم أعداد الرسوم البيانية باستخدام السوراج الجاهزة التالية : SPSS , Harvard , Statistica , Mathematica بالإضافة إلى برنامج الراسم التابع للتوافذ 97 .

يصلح هذا الكتاب كمقرر لطلبة الجامعات في شق التخصصات حيث يدرس الفصول الخمسة الأولى كمقرر وباقي الفصول كمقرر ثاني، كما يصلح لأن يكون مرجعاً لأي باحث . هذا ويمكن للباحثين المستفيدين من هذا الكتاب والذين يعملون في المجالات التطبيقية تخطي براهين النظريات الموجودة في الكتاب .

وأعيراً أود أن أعبر عن شكري الجزيل لكل الذين شجعوني وساعدوني بصورة أو بأخرى على إخراج هذا الكتاب وأخص بالذكر زميلاتي وأعزوني في الله د . سارة السدحان ود . أميرة الضاوي اللتين قلعتا في الكثير من التشجيع والعون سواء في الرسوم البيانية أو في المراجعة اللغوية وأدعو لهما وللجميع بالتوفيق .

كما أتوجه بالشكر لمكتبة الانجلو المصرية التي أتاحت لي الفرصة لنشر هذا العمل العلمي والله ولى التوفيق، وإني أرحب بكل نقد بناء يهدف إلى الأفضل، فالكمال لله وحده .

تقديم

الكتاب الذي بين يدي القارئ عمل مميز يتناول أحد فروع الرياضيات التطبيقية التي تختص بتقديم الطرق والأساليب التي تستخدم في تحليل الظواهر .

استمعت بقراءة هذا الكتاب العلمي الذي صيغت الموضوعات الواردة فيه بطريقة مشوقة سلسلة تجذب القارئ إلى أسلوبه العلمي المفصل ، وهو متنوع الأمثلة مدعم بتمارين شاملة في نهاية كل فصل و مستنداً لعدد كبير من المراجع القيمة ليوسع من مدارك القارئ العربي ولذلك فهو لا غنى عنه لطلاب كليات العلوم بمختلف تخصصاتها كما يمكن أن يخدم طلاب الدراسات العليا للتخصصات الحيوية .

و عبر هذه الصفحات يتضح الجهد الكبير الذي بذلته الدكتورة ثروت محمد عبد المنعم ويوهلها في ذلك ما نالته من غيره المشاركة الفعالة في البحث و التدريس عبر خمس عشر عاما بمحمد الله و توفيقه و التي أتمنى لها مستقبلاً زاهراً.

أ. د الهام شكري

وكيلة معهد الإحصاء - جامعة القاهرة

المحتويات

تقديم

١٧

الفصل الأول : مقدمة

الفصل الثاني : بعض المفاهيم الرياضية

٢٣

١-٢ الفئات

٢٤

٢-٢ عمليات الفئات

٢٦

٣-٢ صيغ الجمع

٢٨

٤-٢ بعض النظريات المفيدة للجمع

٣٠

٥-٢ الدالة

٣٠

تمارين

الفصل الثالث : مقدمة في الاحتمال

٣٥

١-٣ فراغ العينة والأحداث

٣٦

٢-٣ طرق العد

٤٠

٣-٣ الاحتمال

٤١

١-٣-٣ المفهوم القديم (المفهوم الكلاسيكي)

٤٢

٢-٣-٣ مفهوم التكرار النسبي

٤٣

٣-٣-٣ المفهوم الشخصي

٤٣

٤-٣-٣ اخصائص للميزة لقيم الاحتمال

٤٤

٤-٣ بعض قوانين الاحتمال

٤٦

٥-٣ الاحتمال الشرطي

٥٠

٦-٣ الاحتمال الكلي وقاعدة بيز

٥٣

تمارين

الفصل الرابع : التغيرات العشوائية و توزيعاتها الاحتمالية

٦٣

١-٤ المتغير العشوائي

٦٤

٢-٤ التوزيعات الاحتمالية المنفصلة (المتقطعة)

٦٧

٣-٤ التوزيعات الاحتمالية المتصلة (المستمرة)

٦٩

٤-٤ التوقع الرياضي

٧٣	٥-٥ بعض خواص القيم المتوقعة
٧٧	٦-٤ التوزيعات الاحتمالية الثنائية المنفصلة
٨٣	٠ تمارين

الفصل الخامس : عرض ووصف البيانات

٩١	١-٥ المجتمعات و العينات
٩٣	٢-٥ التوزيع التكراري
٩٩	٣-٥ التمثيل البياني
١٠٥	٤-٥ مقاييس الرعة المركزية
١٠٥	١-٤-٥ الوسط الحسابي
١٠٩	٢-٤-٥ الوسيط
١١٢	٣-٤-٥ المنوال
١١٣	٤-٤-٥ الوسط الهندسي
١١٥	٥-٥ الربيعات و المئينات والعشيرات
١١٨	٦-٥ مقاييس التشتت
١١٩	١-٦-٥ المدى و نصف المدى الربيعي
١٢٠	٢-٦-٥ الانحراف المتوسط
١٢١	٣-٦-٥ التباين
١٢٧	٤-٦-٥ مبادئ الاختلاف
١٢٨	٧-٥ الالتواء و العلاقة بين الوسط الحسابي و الوسيط و المنوال
١٢٩	٨-٥ بعض مقاييس الالتواء و التفلطح
١٣٣	٠ تمارين

الفصل السادس : بعض التوزيعات الاحتمالية

١٤٥	١-٦ التوزيع المنتظم
١٤٦	٢-٦ توزيع ذي الحدين
١٥٢	٣-٦ التوزيع الهندسي الزائدي
١٥٦	٤-٦ توزيع بواسون
١٥٨	٥-٦ توزيع ذي الحدين السالب
١٦٠	٦-٦ التوزيع الطبيعي
١٦١	١-٦-٦ التوزيع الطبيعي القياسي

- ١٧٠ ٢-٦-٦ التوزيع الطبيعي كتقريب لذي الحدين
١٧٨ تمارين

الفصل السابع : توزيعات المعاينة

- ١٩٣ ١-٧ مقدمة
١٩٤ ٢-٧ توزيعات المعاينة الطبيعية
١٩٤ ٣-٧ توزيعات المعاينة للمتوسط
٢٠٥ ٤-٧ التوزيعات العينة للفرق بين متوسطي مجتمعين
٢١٠ ٥-٧ التوزيعات العينة للنسب
٢١٧ ٦-٧ توزيع ١
٢٢٥ ٧-٧ توزيع مربع كاي
٢٢٧ ٨-٧ توزيع F
٢٣٠ تمارين

الفصل الثامن : فترات الثقة

- ٢٤١ ١-٨ مقدمة
٢٤٣ ٢-٨ فترة ثقة لمتوسط المجتمع μ
٢٤٩ ٣-٨ فترة ثقة للفرق بين متوسطي مجتمعين $\mu_1 - \mu_2$
٢٥٩ ٤-٨ فترة ثقة للنسبة
٢٦٢ ٥-٨ فترة ثقة للفرق بين نسبتين
٢٦٤ ٦-٨ فترة ثقة للتباين
٢٦٧ ٧-٨ فترة ثقة لنسبة تباينين
٢٦٩ تمارين

الفصل التاسع : اختبارات الفروض

- ٢٧٩ ١-٩ الفروض الإحصائية
٢٧٩ ٢-٩ الخطأ من النوع الأول و الخطأ من النوع الثاني
٢٨٥ ٣-٩ اختبارات من جانب واحد أو من جانبيين
٢٨٦ ٤-٩ اختبارات حول متوسط المجتمع μ
٢٩٢ ٥-٩ اختبارات حول تباين المجتمع σ^2
٢٩٤ ٦-٩ اختبارات تخص تباين مجتمعين

٢٩٦	٧-٩ اختبارات فحص المتوسطات
٣٠٣	٨-٩ اختبارات ، للأزواج
٣٠٥	٩-٩ اختبارات فحص نسبة مجتمع
٣٠٦	١٠-٩ اختبارات فحص الفرق بين نسبي مجتمعين
٣٠٧	تمارين

الفصل العاشر : الانحدار و الارتباط

٣٣١	١-١٠ الانحدار
٣٣١	٢-١٠ الانحدار الخطي البسيط
٣٣٢	١-٢-١٠ شكل الانتشار
٣٣٣	٢-٢-١٠ بناء نموذج الانحدار البسيط
٣٣٤	٣-٢-١٠ طريقة المربعات الصغرى
٣٣٦	٤-٢-١٠ تحليل الانحدار
٣٤٠	٥-٢-١٠ تقدير σ^2
٣٤١	٦-٢-١٠ معامل التحديد البسيط
٣٤٣	٧-٢-١٠ تقدير المعلمتين β_0, β_1
٣٤٧	٨-٢-١٠ التنبؤ
٣٥٠	٩-٢-١٠ اعتبار خطية الانحدار
٣٥٤	٣-١٠ بعض نماذج الانحدار الغير خطي
٣٥٤	١-٣-١٠ النموذج الاسي
٣٥٦	٢-٣-١٠ نموذج القوي
٣٥٨	٤-١٠ معامل الارتباط الخطي البسيط
٣٦٤	٥-١٠ الانحدار الخطي المتعدد
٣٦٤	١-٥-١٠ طريقة المربعات الصغرى
٣٦٧	٢-٥-١٠ تحليل الانحدار
٣٦٩	٣-٥-١٠ معامل التحديد المتعدد
٣٧٠	٤-٥-١٠ الارتباط المتعدد و الجزئي
٣٧٢	٦-١٠ الانحدار من الدرجة الثانية
٣٧٤	تمارين

الفصل الحادي عشر : تحليل التباين

٣٩٥	١-١١ مقدمة
٣٩٦	٢-١١ التصنيف الأحادي
٤٠٣	٣-١١ اختبار تجانس عدة تباينات
٤٠٤	٤-١١ اختبار دايكن للمدى المتعدد
٤٠٨	٥-١١ التصنيف الثنائي ، مشاهدة واحدة لكل خلية
٤١٦	٦-١١ التصنيف الثنائي ، عدة مشاهدات لكل خلية
٤٢٤	تمارين

الفصل الثاني محشر: الاختبارات اللامعلمية

٤٤٩	١-١٢ مقدمة
٤٥٠	٢-١٢ اختبار مربع كاي لجودة التوفيق
٤٥٦	٣-١٢ اختبار مربع كاي للاستقلال
٤٦٢	٤-١٢ اختبار مربع كاي للتجانس
٤٦٥	٥-١٢ اختبار خاص بالاعتدال
٤٦٦	٦-١٢ اختبار الإشارة لعينة واحدة
٤٦٩	٧-١٢ اختبار الإشارة لميتين مرتبطتين (عينة مزدوجة)
٤٧١	٨-١٢ اختبار إشارة الترتيب
٤٧٤	٩-١٢ اختبار Mann-Whitney-Wilcoxon
٤٧٦	١٠-١٢ اختبار Kruskal-Wallis
٤٨٠	١١-١٢ اختبار الدورات
٤٨٣	١٢-١٢ معامل ارتباط سيرومان للترتيب
٤٨٧	تمارين

المراجع

٥١٧

الملاحق

الفصل الأول

مقدمه

Introduction

يهتم الإحصائي ، أساسا ، بالنواتج التي يحصل عليها الباحثون من إجراء الأبحاث العلمية . فعلى سبيل المثال قد يهتم بعدد الحوادث التي تقع في تقاطع صا، عدد الأفراد في الأسرة، عدد البكتريا في لتر من الماء النقي، عدد المصابين بالسرطان، كمية الحديد التي تستهلكها الإناث البالغات، عدد المصابين بتسوس الأسنان، أطوال مجموعة من الأشخاص... الخ ، أي أنه يهتم بقراءات (ملاحظات) قابلة للعد أو للقياس . سوف نشير إلى المعلومات *information* المسجلة في صورتها الأصلية التي جمعت بما بالبيانات الخام *raw data* . يهتم الإحصائي بالطرق المختلفة لوصف كمية كبيرة من البيانات الخام وذلك بغرض الوصول إلى معلومات لم تكن معلومة له . أو قد يهتم بالوصول إلى استنتاجات أو اتخاذ قرارات عن فئة كبيرة من البيانات وذلك بالاعتماد على فئة جزئية منها .

— يهتم علم الإحصاء بالطرق المستخدمة في جمع و عرض و وصف و تحليل و تفسير البيانات . كل المشاكل التي تستخدم الطرق الإحصائية يمكن أن تنتمي إلى مجال الإحصاء الوصفي *descriptive statistics* أو الإحصاء الاستقرائي *inductive statistics* ، أي معالجة للبيانات تؤدي إلى تنبؤات أو استدلالات تخص مجموعة كبيرة من البيانات تجعلنا في مجال الإحصاء الاستقرائي . ومن ناحية أخرى إذا كان اهتمامنا منحصرا في البيانات التي في حوزتنا ولا توجد أي محاولات لتعميمها إلى فئة أكبر من البيانات، فإننا نكون في مجال الإحصاء الوصفي . ويمكن توضيح الفرق بين الإحصاء الوصفي والإحصاء الاستقرائي من المثال التالي : فإذا كان اهتمامنا بفئة من القيم الممثلة لكمية الصادرات لسلعة ما في منطقة ما خلال العشرين سنة الماضية، فإن أي قيمة تصف البيانات، مثل متوسط كمية الصادرات في العشرين سنة الماضية ، هي قيمة في مجال الإحصاء الوصفي . في هذه الحالة لم نحاول أن نذكر أي شئ عن كمية الصادرات لأي سنة أخرى غير السنوات العشرين والتي حصلنا منها على المعلومات . فإذا كان متوسط كمية الصادرات في هذه المنطقة في العشرين سنة الماضية 1.34 طن واستطعنا أن نضع جملة نص على أن كمية الصادرات في السنة القادمة سوف تتراوح بين 1.2 إلى 1.7 طن، في هذه الحالة نكون قد عممنا ووضعنا أنفسنا في مجال الإحصاء الاستقرائي .

التعميم في الإحصاء الاستقرائي يخضع للأحداث غير المؤكدة لأننا نحتم فقط بمعلومات جزئية ثم الحصول عليها من فئة جزئية من البيانات موضع الدراسة . للتعامل مع الأحداث غير المؤكدة فإن فهمنا لنظرية الاحتمالات يكون ضروريا . في هذا الكتاب سوف نقدم بعض المفاهيم الأساسية في نظرية الاحتمالات وذلك في الفصل الثالث والرابع . وبما أنه يمكن عرض الاحتمالات بصورة أفضل وباستخدام صيغ الفئات لذلك سوف نقدم بعض الأساسيات عن

الفئات وخصوصها في الفصل الثاني . أما بالنسبة للرياضيات المطلوبة لمقرر الإحصاء فلا تصدى بعض الصيغ الأساسية في مقرر الجبر الذي يدرس في الجامعات . بعض الصيغ الرياضية المفيدة في مجال الإحصاء سوف نتناولها في الفصل الثاني .

يتناول هذا الكتاب الأسس العامة لمبادئ الإحصاء والاحتمالات . تتناول الفصول الأولى من الكتاب بعض الصيغ الضرورية لدراسة الإحصاء والاحتمالات، والتغيرات العشوائية، وبعض التوزيعات الاحتمالية، وتطبيقاتها، وطرق عرض ووصف البيانات وما يتعلق بها من مقاييس الالتواء والتضلع . ويتطرق الكتاب في الفصول الأخيرة إلى بعض توزيعات المعانسة واستخدامها في إيجاد فترات الثقة واختبارات الفروض وتحليل التباين والارتباط والإمحدار . أما الفصل الأخير فيهتم بالاختبارات اللامعلمية .

يقدم هذا الكتاب المعلومات بالوسيلة التي تفيد كل التخصصات سواء في الطب، الزراعة، الأعمال . . . الخ . فالطرق الأساسية لجمع وعرض وتحليل البيانات واحدة بصرف النظر عن مجال التطبيق . فعلى سبيل المثال قد يقوم الكيميائي بإجراء تجربة ما مستخدماً أربع طرق بفرض قياس كمية منتج ما، موضع الدراسة، ثم تحليل النتائج باستخدام الطرق الإحصائية . هذه الطرق نفسها، يمكن استخدامها لتحليل البيانات التي تمثل عدد الوحدات المنتجة الناتجة من استخدام أربع آلات مختلفة أو لتحليل البيانات التي يتم الحصول عليها من قياس كمية المحصول عند اختبار أربعة أسمدة مختلفة . كثير من الطرق المصممة للتطبيقات الزراعية أثبتت كفاءتها في مجالات أخرى .

في الحقيقة يعتبر الإحصاء أداة قوية جداً وضرورية إذا ما استخدمت بطريقة صحيحة . ولذلك لابد من تطبيق الطرق الصحيحة والأكثر كفاءة للظروف المعطاة وذلك من أجل الحصول على أقصى معلومات من البيانات المتوفرة . الطرق المستخدمة لتحليل فئة من البيانات تعتمد بدرجة كبيرة على الطريقة المستخدمة في جمع البيانات . ولهذا السبب يكون من المفصل استشارة الإحصائي من بداية التخطيط للبحث، حتى الوصول إلى النتائج وتحليلها وتفسيرها .

الفصل الثاني

بعض المفاهيم الرياضية

Some Mathematical Concepts

يتناول هذا الفصل بعض المفاهيم الأساسية في الرياضيات والضرورية للدراسة الإحصاء .

١- الفئات Sets

الفئة هي مجموعة من الأشياء المعرفة تماما . على سبيل المثال الأتجار في أفريقيا، الأعداد الصحيحة الموجبة $1, 2, 3, \dots$ ، الحروف الأبجدية من أ إلى ي كلها تمثل فئات . الأشياء التي تكون تسمى عناصر **elements** أو أعضاء **members** في الفئة . عادة يمكن تمثيل ف الإنجليزية الكبيرة مثل A, B, C, Y . هناك طريقتان لوصف الفئة، إذا احتوت الفئة على عدد محدود من العناصر بحيث يمكن عمل قائمة بهذه العناصر، فعلى سبيل المثال الفئة A والتي تتكون من العناصر $2, 5, 6, 7$ يمكن كتابتها على الشكل $A = \{2, 5, 6, 7\}$ أو الفئة B ، والتي تمثل نوايج إلقاء زهرة نرد يمكن كتابتها على الشكل $B = \{1, 2, 3, 4, 5, 6\}$. يمكن وصف الفئة بجملة **statement** أو قاعدة **rule** فعلى سبيل المثال إذا كانت Y تمثل الفئة من كل الأشخاص في العالم وإذا كان y عنصر اختياري في Y ، فإنه يمكن كتابة الفئة Y على الشكل :

$$Y = \{y \mid y \text{ is a person living in the earth} \}.$$

ويعتمد وصف الفئة، سواء بقائمة أو قاعدة على نوع المشكلة موضع الدراسة . على سبيل المثال يكون من الصعوبة وضع قائمة بعناصر الفئة من الأزهار الحمراء في العالم . ومن ناحية أخرى لا توجد قاعدة بسيطة لوصف الفئة :

$$D = \{\text{family, book, flower}\}.$$

باستخدام صيغ الفئة فإن الرمز \in يعني " أنه عضو في " أو " ينتمي إلى " والرمز \notin يعني " ليس عضو في " أو " لا ينتمي إلى " . إذا كان x عنصر في الفئة A و y ليس عنصر في A فإن $x \in A, y \notin A$.

مثال (١-٢) ليكن $A = \{2, 4, 6, 8\}$, $B = \{x \mid x \text{ is an integer divisible by } 7\}$ فإن $8 \in A$, $3 \notin A$, $49 \in B$

تعريف : تساوى فئتين إذا احتوت الاثنتان بالضبط على نفس العناصر .

إذا كانت الفئة A تساوى أو تماثل الفئة B ، فإن كل عنصر ينتمي إلى A ينتمي إلى B ، وكل عنصر ينتمي إلى B ينتمي إلى A . وسوف نرمز لهذا التساوي بكتابة $A = B$. في بعض الأحيان، إذا كانت إحدى الفئتين A أو B تحتوي على عنصر واحد على الأقل لا ينتمي

إلى الاثنتين، فلنأخذ نقول أن الفئتين غير متساويتين، وفي هذه الحالة نكتب $A \neq B$.

مثال (٢-٢) ليكن $A = \{1, 4, 5\}$, $B = \{4, 6, 9\}$, $C = \{1, 5, 4\}$ فإن $A \neq B$:-

$$A = C, B \neq C.$$

ويجب أن نتذكر أن الفئات لا تتغير عندما نغير ترتيب العناصر.

تعريف : الفئة الخالية **empty** أو فئة العدم **null set** هي الفئة التي لا تحوى على أى عناصر ويرمز لها بالرمز ϕ .

إذا كانت :

$$A = \{x | x \text{ is a letter before A in the alphabet}\}$$

و

$$B = \{x | x^2 = 4, x \text{ is an odd number}\}$$

فإن A و B فئتان خاليتان.

لتكن الفئة $A = \{1,4\}$, $B = \{1,4,7,8,9\}$. يلاحظ أن كل عنصر في الفئة A

هو أيضا عنصر في الفئة B . الفئة A تسمى فئة جزئية **subset** من B وتكتب $A \subset B$.

تعريف : إذا كان كل عنصر في الفئة A هو عنصر في الفئة B فإن A تسمى فئة جزئية من B .

بناء على ذلك، فإن أى فئة تعتبر فئة جزئية من نفسها.

في كثير من المناقشات كل الفئات تعتبر فئات جزئية من فئة واحدة خاصة، هذه الفئة

تسمى الفئة الشاملة ويرمز لها بالرمز U .

مثال (٢-٣) كل الفئات الجزئية من الفئة الشاملة $U = \{4,5,6\}$ هي :-

$\{4\}, \{5\}, \{6\}, \{4,5\}, \{4,6\}, \{5,6\}, \{4,5,6\}, \{\phi\}$ أى أن عدد الفئات الجزئية من الفئة

الشاملة التي عدد عناصرها n هو 2^n من الفئات الجزئية.

Sets Operations

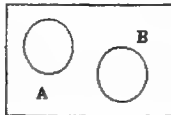
(٢-٢) عمليات الفئات

سوف نعرض بعض العمليات على الفئات والتي تنتج فئات جديدة. هذه الفئات

الجديدة تعتبر فئات جزئية من الفئة الشاملة. العلاقة بين الفئات الجزئية والفئة الشاملة يمكن

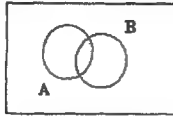
توضيحها بما يسمى شكل **Venn diagram**. تمثل الفئة الشاملة بمسقطيل و الفئات

الجزئية بدوائر داخل المستطيل كما في شكل (٢-١)



شكل (٢-١)

سوف نرمز لتقاطع القتين A , B بالرمز $A \cap B$. العناصر في $A \cap B$ لابد أن ينتمي إلى كل من A , B . يوضح الشكل (٢-٢) الفئة $A \cap B$ بالجزء المظلل.



شكل (٢-٢)

مثال (٢-٤) إذا كان $A = \{1,7,8,9,10\}$, $B = \{2,8,9,10,11\}$

فإن $A \cap B = \{8,9,10\}$

مثال (٢-٥) إذا كان :-

$A = \{x | x \text{ is an integer and } 1 \leq x \leq 6\}$,

$B = \{y | y \text{ is an integer greater than } 4\}$

فإن $B = \{5,6,7,8,\dots\}$, $A \cap B = \{5,6\}$

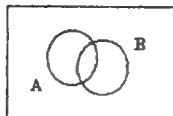
تعريف : إذا كان $A \cap B = \emptyset$ ، يقال للقتين A , B أنهما منفصلتان disjoint ، أي لا يوجد أي عناصر مشتركة بينهما كما في شكل (٢-٣) .



شكل (٢-٣)

تعريف : الاتحاد بين قتين A , B هو الفئة من العناصر التي تنتمي إلى A أو B أو كلاهما .

سوف نرمز للاتحاد بين A و B بالرمز $A \cup B$. يوضح الشكل (٢-٤) الفئة $A \cup B$ بالجزء المظلل.

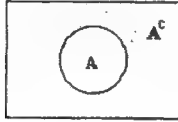


شكل (٢-٤)

مثال (٢-٦) إذا كان :-

$$A \cup B = \{1,5,7,8,9,10,12\} \text{ فإن } A = \{1,5,7,8\} , B = \{8,9,10,12\}$$

إذا كانت A فئة جزئية من الفئة الشاملة U ، فإن مكمل الفئة A هو الفئة من العناصر في U والغير موجودة في A . سوف نرمز لمكمل الفئة A بالرمز A^c . ويوضح شكل (٢-٥) الفئة المكملية بالجزء المظلل .



شكل (٢-٥)

مثال (٢-٧) بفرض أن :-

$$A^c = \{8,9,10,12\} \text{ فإن } U = \{1,5,7,8,9,10,12\} , A = \{1,5,7\}$$

هناك عدة نتائج من التعاريف السابقة مثل :-

$$A \cap \phi = \phi$$

$$A \cup \phi = A$$

$$A \cap A^c = \phi$$

$$A \cup A^c = U$$

$$U^c = \phi$$

$$\phi^c = U$$

$$(A^c)^c = A$$

Summation Notation

(٢-٣) صيغ الجمع

سوف نحتاج في التحليل الإحصائي للبيانات إلى جمع مجموعة من الأعداد . إذا كانت x_i تمثل أى قيمة من القيم التالية التي عددها n : x_1, x_2, \dots, x_n والناتجة للمتغير X . الحرف i في x_i ، الذي يأخذ أى من الأرقام $1, 2, \dots, n$ ، يسمى الدليل subscript أو index . من الواضح أنه يمكن استخدام أى حرف غير i مثل s, p, q, j, k الصيغة

$\sum_{i=1}^n x_i$ تستخدم لتمثيل جمع كل قيم x_i من $i=1$ إلى $i=n$ • الرمز \sum يسمى سيجما • وعلى ذلك فإن :

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n.$$

للتبسيط عادة يكتب المجموع بأشكال كثيرة مثل $\sum x$ أو $\sum x_i$ أو $\sum x_i$ إذا كان :-

$$\text{فإن } x_1 = 2, x_2 = 3, x_3 = 5, x_4 = 1, x_5 = 2$$

$$\sum_{i=1}^5 x_i^2 = (2)^2 + (3)^2 + (5)^2 + (1)^2 + (2)^2 = 43.$$

مثال (٢-٨) إذا كان $x_1 = 2, x_2 = 5, x_3 = 6, x_4 = 3, x_5 = 4$ أوجد :-

$$\sum_{i=3}^5 4x_i \quad (ب) \quad \sum_{i=2}^5 (x_i - 1) \quad (أ)$$

الحل •

$$\sum_{i=2}^5 (x_i - 1) = (5-1) + (6-1) + (3-1) + (4-1) \quad (أ)$$

$$= 4 + 5 + 2 + 3 = 14.$$

$$\sum_{i=3}^5 4x_i = (4)(6) + (4)(3) + (4)(4) = 52. \quad (ب)$$

مثال (٢-٩) إذا كان $x_1 = 5, x_2 = 4, x_3 = 3, x_4 = 1$ أوجد :-

$$\sum_{i=1}^3 x_i - 1 \quad (جـ) \quad \sum_{i=1}^3 (c - i + 2) \quad (ب) \quad \sum_{i=1}^4 (x_i - c) \quad (أ)$$

الحل •

(أ)

$$\sum_{i=1}^4 (x_i - c) = (x_1 - c) + (x_2 - c) + (x_3 - c) + (x_4 - c)$$

$$= (5 - c) + (4 - c) + (3 - c) + (1 - c) = 13 - 4c.$$

(ب)

$$\sum_{i=1}^3 (c - i + 2) = (c - 1 + 2) + (c - 2 + 2) + (c - 3 + 2)$$

$$= 3c.$$

(جـ)

$$\sum_{i=1}^3 x_i - 1 = (5 + 4 + 3) - 1 = 11.$$

Useful Theorems Relating to Sums (٢-٤) بعض النظريات المفيدة للجمع

نظرية (٢-١) إذا كان c عدد حقيقي فإن :-

$$\sum_{i=1}^n c = nc.$$

البرهان :-

$$\sum_{i=1}^n c = c + c + \dots + c.$$

حيث أن c تتكرر n من المرات فإن :-

$$\sum_{i=1}^n c = nc.$$

مثال (٢-١٠)

$$\sum_{i=1}^5 4c = 5(4c) = 20c.$$

مثال (٢-١١)

$$\sum_{i=1}^4 (5c - 3) = 4(5c - 3).$$

نظرية (٢-٢) إذا كان c عدد حقيقي فإن :-

$$\sum_{i=1}^n cx_i = c \sum_{i=1}^n x_i.$$

البرهان :-

$$\begin{aligned} \sum_{i=1}^n cx_i &= cx_1 + cx_2 + \dots + cx_n \\ &= c(x_1 + x_2 + \dots + x_n) \\ &= c \sum_{i=1}^n x_i. \end{aligned}$$

مثال (٢-١٢) للبيانات في مثال (٢-٩) :-

$$\sum_{i=2}^4 7x_i = 7(4 + 3 + 1) = 56.$$

نظرية (٣-٢)

$$\sum_{i=1}^n (x_i + y_i + z_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i + \sum_{i=1}^n z_i.$$

البرهان :-

$$\sum_{i=1}^n (x_i + y_i + z_i) = x_1 + y_1 + z_1 + x_2 + y_2 + z_2 + x_3 + y_3 + z_3$$

$$+ \dots + x_n + y_n + z_n$$

وبإعادة التجميع :-

$$\sum_{i=1}^n (x_i + y_i + z_i) = (x_1 + x_2 + \dots + x_n) + (y_1 + y_2 + \dots + y_n)$$

$$+ (z_1 + z_2 + \dots + z_n) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i + \sum_{i=1}^n z_i.$$

مثال (١٣-٢) للبيانات في مثال (٩-٢) :-

$$\sum_{i=1}^4 (x_i^2 + cx_i + 6) = \sum_{i=1}^4 x_i^2 + \sum_{i=1}^4 cx_i + \sum_{i=1}^4 6$$

$$= \sum_{i=1}^4 x_i^2 + c \sum_{i=1}^4 x_i + 4(6)$$

$$= (5)^2 + (4)^2 + (3)^2 + (1)^2 + c(5 + 4 + 3 + 1) + 24$$

$$= 51 + 13c + 24 = 75 + 13c.$$

مثال (١٤-٢)

$$\sum_{i=1}^3 (c^2 + 7i) = \sum_{i=1}^3 c^2 + \sum_{i=1}^3 7i$$

$$= 3c^2 + 7 \sum_{i=1}^3 i$$

$$= 3c^2 + 7(1 + 2 + 3)$$

$$= 3c^2 + 42.$$

(٥-٢) الدالة Function

تعريف : بفرض أن X, Y فئتان غير خاليتين القاعدة التي تعين لكل عنصر $x \in X$

عنصر وحيد $y \in Y$ تسمى دالة من X إلى Y .

العنصر الوحيد y المعين بدالة ما والمناظر لعنصر محدد x يقال له قيمة الدالة عند x ويرمز للدالة بالرمز f, p, y, h, \dots الخ .

مثال (١٥-٢) إذا كانت X, Y فئتين من الأعداد حقيقية وكان $y = x + 3$ فإن y دالة في x . عندما $x = 1$ فإن $y = 4$ وعندما $x = -3$ فإن $y = 0$ وهكذا .

بفرض أن f دالة ما معطاة فإن الفئة X التي تعين الدالة f لكل عنصر من عناصرها عنصرا وحيدا $y \in Y$ يقال لها مجال domain الدالة . الفئة التي عناصرها العناصر المناظرة $y \in Y$ المعينة بالدالة f يقال لها مدى range الدالة .

إذا كانت الدالة f تعين قيمة y في مداها لعنصر x في مجالها فإننا نكتب $y = f(x)$ وتسمى y قيمة f عند x . بفرض أن $y = 5x + 3$ وباستخدام صيغة الدالة فإننا يمكن أن نكتب $f(x) = 5x + 3$ وعلى ذلك $f(6) = 5(6) + 3 = 33$.

مثال (١٦-٢) إذا كانت $g(x) = x^2$ المطلوب إيجاد (١) $g(2)$ (ب) المجال والمدى للدالة $g(x)$.

الحل . (١) $g(2) = 2^2 = 4$ (ب) المجال للدالة $g(x)$ هو جميع الأعداد الحقيقية والمسمى جميع الأعداد الحقيقية الغير سالبة .

مثال (١٧-٢) إذا كانت $p(x) = \frac{1}{x^2}$ المطلوب إيجاد (١) $p(3)$ (ب) المجال والمسمى للدالة $p(x)$.

الحل . (١) $p(3) = \frac{1}{3^2} = \frac{1}{9}$ (ب) المجال للدالة $p(x)$ هو جميع الأعداد الحقيقية ما عدا $x = 0$ والمسمى جميع الأعداد الحقيقية الغير سالبة ما عدا $y = 0$.

تمارين

١ - أكتب عناصر الفئات التالية :-

(١) الفئة من الأعداد الصحيحة بين 9 , 13 .

(ب) الفئة من الأعداد الصحيحة أقل من 9 .

(جـ) الفئة من الأعداد الصحيحة بين 1 و 20 القابلة للقسمة على 3.

٢ - أكتب عناصر كل من الفئات التالية :-

$$A = \{x | x^2 + x - 6 = 0\}$$

(ب) فئة نواتج إلقاء زهرة نرد وعملة في نفس الوقت .

$$B = \{x | 2x - 4 = 0 \text{ and } x > 5\}$$

٣ - بفرض أن $A = \{x | 3x = 6\}$ هل الجملة $A = 2$ صحيحة ؟

٤ - لتكن $E = \{3, 5, 6, 8, 9\}$ أى من الجمل التالية صحيح وأي خطأ ؟

$$5 \in E \quad (ب) \quad 7 \in E \quad (جـ) \quad 9 \notin E \quad (د) \quad \phi \subset E$$

٥ - أى الفئات التالية تتساوى ؟

$$\{t, s, r\}, \{r, s, t\}, \{t, r, s\}, \{s, r, t\}.$$

٦ - أكتب عناصر كل من الفئات التالية :

$$A = \{x | x \text{ is an integer and } 9 \leq x \leq 13\}$$

$$C = \{x | x + 2 = 0\}$$

$$D = \{y | y^2 + 3y = 28\}$$

٧ - لتكن $U = \{x, y, 6, z\}$ أكتب كل الفئات الجزئية من U .

٨ - أكتب كل الفئات الجزئية من $U = \{\text{man, woman, baby, home}\}$

٩ - لتكن $U = \{1, 2, \dots, 9\}$, $A = \{1, 2, 3, 4\}$, $B = \{2, 4, 6, 8\}$, $C = \{3, 4, 5, 6\}$

أوجد :- (أ) A^c (ب) $A \cap B$ (جـ) $(A \cap C)^c$ (د) $A \cup B$

١٠ - لتكن $U = \{10, 11, 12, 13, 14, 15\}$, $A = \{10, 11, 12, 13\}$, $B = \{12, 13, 14\}$

أوجد :- (أ) $A \cup B$ (ب) $A \cap B$ (جـ) A^c

١١ - إذا كانت :-

$$A = \{x | x \text{ is an integer and } x > 17\},$$

$$B = \{x | x \text{ is an integer and } 12 < x < 20\}$$

أوجد $A \cap B$

١٢ - إذا كانت :-

$$A = \{x | x^2 - 7x = 8\},$$

$$B = \{x | 2x - 16 = 0\}$$

أوجد $A \cap B$

١٣ - إذا كانت :-

$$X = \{x | x^2 + x - 12 = 0\},$$

$$Y = \{y | y^2 = 4\}$$

$$X \cap Y, X \cup Y \text{ أوجد}$$

١٤ - إذا كانت :

$$A = \{x | x^2 + x - 12 = 0\},$$

$$B = \{x | 2x - 6 = 0\}$$

$$A \cap B \text{ أوجد}$$

١٥ - إذا كانت :-

$$\text{أوجد :- } x_1 = 2, x_2 = 5, x_3 = 6, x_4 = 3, x_5 = 4, x_6 = 9, x_7 = 11$$

$$\sum_{i=1}^4 (x_i + c)^2 \quad (\text{جـ}) \quad \sum_{i=1}^5 (x_i + c) + 2 \quad (\text{ب}) \quad \sum_{i=1}^7 (5x_i + 3) \quad (١)$$

١٦ - إذا كانت :-

$$\text{أوجد :- } x_1 = 8, x_2 = 9, x_3 = 6, x_4 = 3, x_5 = 4, x_6 = 7$$

$$\sum_{i=1}^6 (x_i^3 + 3i) \quad (\text{جـ}) \quad \sum_{i=1}^6 (x_i + x_i c^2)^3 \quad (\text{ب}) \quad \sum_{i=1}^6 (x_i^2 + x_i + 3) \quad (١)$$

$$\text{أوجد :- } x_1 = 7, x_2 = 5, x_3 = 2, x_4 = 10, x_5 = 8$$

$$\sum_{i=1}^5 (x_i + i)^2 \quad (\text{جـ}) \quad \sum_{i=1}^5 (x_i + x_i^2 c)^2 \quad (\text{ب}) \quad \sum_{i=1}^5 (x_i^2 + 6x_i) \quad (١)$$

$$\text{أوجد :- } f(x) = 3x^2 + 4x + 2$$

$$f(0) \quad (١) \quad f(-0.5) \quad (\text{ب}) \quad f(3) \quad (\text{جـ})$$

$$\text{١٩ - يعطى حجم مستعمرة حشرية عند لحظة زمنية } t \text{ (الزمن قيس بالأيام) بالعلاقة :-}$$

$$f(t) = 5000 - \frac{1000}{1 + t^2}$$

$$\text{أوجد :- } f(0) \quad (١) \quad f(2) \quad (\text{ب}) \quad f(5) \quad (\text{جـ})$$

$$\text{٢٠ - يعطى حجم مزرعة بكتريا عند لحظة زمنية } t \text{ (الزمن قيس بالساعات) بالعلاقة :-}$$

$$f(t) = 2000 + 1000t - 1000t^2$$

$$\text{أوجد :- } f(0) \quad (١) \quad f(2) \quad (\text{ب}) \quad f(5) \quad (\text{جـ})$$

الفصل الثالث

مقدمة في الاحتمال

Introduction to Probability

(١-٣) فراغ العينة والأحداث Sample Space and events

تُجرى الأبحاث في مجالات كثيرة، ففي مجال الطب قد يهتم باحث بدراسة تأثير دواء معين على الشفاء من مرض ما، وفي مجال الاقتصاد قد يهتم باحث بدراسة أسعار ثلاث سلع مختلفة في فترات زمنية مختلفة، وفي مجال الزراعة قد يهتم باحث بدراسة تأثير سماد كيميائي على كمية الحصول. الطريق الوحيد للباحث للحصول على معلومات عن الظاهرة موضع الدراسة هو إجراء تجربة **experiment** وهي أى إجراء نحصل به على بيان (مشاهدة) سواء في الطبيعة أو في العمل وهذا البيان قد يكون رقمي أو وصفي. كل مثال من الأمثلة التالية يمثل تجربة :

(أ) تسجيل درجة طالب.

(ب) قياس كمية المطر في يوم ما.

(جـ) فحص مصباح كهربائي وتسجيل نموه.

(د) ملاحظة وحدة منتجة وتدوين حالتها : سليمة أو تالفة.

(هـ) إلقاء عملة وملاحظة الوجه الذي يظهر.

نجد في معظم الحالات أن نتيجة التجربة تعتمد على عوامل الصدفة (عوامل خارجة عن إرادة الباحث أى في علم الله) ولا يمكن التنبؤ بها بشيء من التأكد، ولكن يمكن وصف فئة كل النتائج الممكنة لها قبل إجرائها.

تعريف : الفئة التي عناصرها تمثل جميع النواتج الممكنة لتجربة تسمى فراغ العينة.

مثال (١-٣) عند إلقاء زهرة نرد فإن فراغ العينة هو:-

$$S = \{1,2,3,4,5,6\}.$$

تعريف : يسمى أى عنصر في فراغ العينة نقطة عينة **sample point**.

مثال (٢-٣) عند إلقاء عملة ثلاث مرات فإن فراغ العينة هو :-

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

حيث H ترمز إلى ظهور صورة **head** و T ترمز إلى ظهور كتابة **tail**. فراغ العينة في هذه التجربة يتكون من 8 نقاط عينة.

تعريف : الحادثة **event** هي أى فئة جزئية من فراغ العينة.

في المثال (٢-٣) تمثل $A = \{HHH\}$ حادثة "ظهور ثلاث صور"، أيضاً $B = \{HHH, HHT, HTH, THH\}$ تمثل حادثة "ظهور صورتين على الأقل" ونعتبر

A, B فئتان جزئيتان من فراغ العينة.

تعريف : إذا كانت الفئة الجزئية تحتوي على عنصر واحد فقط تسمى حادثة بسيطة simple event . أما الحادثة المركبة compound event فهي التي تنتج من اتحاد أحداث بسيطة. مثال (٣-٣) عند إلقاء عملتين مرة واحدة فإن فراغ العينة هو $S = \{HH, HT, TH, TT\}$. تمثل الحادثة $A = \{HH\}$ " ظهور صورتين " حادثة بسيطة بينما تمثل الحادثة $B = \{HH, HT, TH\}$ " ظهور صورة واحدة على الأقل " حادثة مركبة. نقول أن الحادثة وقعت إذا ظهر أحد عناصرها عند إجراء التجربة. تمثل الفئة الخالية ϕ الحادثة المستحيلة الحدوث كما يمثل فراغ العينة S الحادثة المؤكدة الحدوث.

تعريف : يقال أن A, B حادثتان مانعتان (متنافيتان) exclusive events إذا كان لوسوع إحداها يمنع وقوع الآخر وفي هذه الحالة فإن $P(A \cap B) = \phi$. فمثلا عند إلقاء عملة مرة واحدة فإن ظهور الصورة يمنع ظهور الكتابة وبالتالي فإن الحادثة التي تمثل ظهور الصورة والحادثة التي تمثل ظهور الكتابة حادثتان مانعتان. مثال (٤-٣) عند إلقاء زهرة نرد مرة واحدة، يفرض أن الحادثة $A = \{2,4,6\}$ ظهور رقم زوجي والحادثة $B = \{1,3,5\}$ ظهور رقم فردي فإن $P(A \cap B) = \phi$. وعلى ذلك فإن A و B حادثتان مانعتان.

Counting Methods

(٢-٣) طرق العد

يعتبر عنصر الصدفة المرتبط بظهور بعض الأحداث من المشاكل التي يقابلها الإحصائي ويحاول تقديرها عند إجراء التجربة . تنتمي هذه المشاكل إلى فرع الاحتمال والسذني سوف نتناوله في البند التالي. في كثير من الحالات نكون قادرين على حل مشكلة الاحتمال عن طريق عد النواتج التي تنتمي إلى الحادثة محل السؤال وأيضاً العدد الكلي لنواتج التجربة . ولكن لبعض التجارب قد يكون عدد النواتج كبير جداً، وبالتالي قد يكون عمل قائمة تضمهم جميعاً مهمة طويلة وصعبة. القاعدة الأساسية للعد في النظرية الآتية.

نظرية (١-٣) إذا أمكن إجراء عملية ما بطرق عددها n_1 وإذا أمكن إجراء عملية ثانية بطرق عددها n_2 و ... وإذا أمكن إجراء عملية k بطرق عددها n_k ، فإنه يمكن إجراء هذه العمليات معاً بطرق عددها $n_1 \cdot n_2 \cdot \dots \cdot n_k$.

مثال (٥-٣) توجد ثلاث طرق بين A, B وتوجد أربع طرق بين B, C بكم طريقة يمكن لسائق أن يصل من A إلى C .

الحل . عدد الطرق الممكنة للوصول من A إلى B هو $n_1 = 3$ وعدد الطرق الممكنة للوصول من B إلى C هو $n_2 = 4$ وعلى ذلك عدد الطرق الممكنة للوصول من A إلى C هو :-

$$n_1 \cdot n_2 = 3 \cdot 4 = 12.$$

مثال (٣-٦) يقدم مطعم 5 أصناف من اللحم و 3 أصناف من الحساء و 3 أصناف من السلطة و 4 أصناف من العصير . كم عدد الوجبات المختلفة التي يمكن تقديمها إذا علم أن الوجبة الواحدة تتكون من لحم وحساء وسلطة وعصير ؟

الحل . عدد الوجبات المختلفة التي يمكن تقديمها هي :-

$$n_1 \cdot n_2 \cdot n_3 \cdot n_4 = 5 \cdot 3 \cdot 3 \cdot 4 = 180.$$

عادة يكون الاهتمام بفراغ العينة الذي عناصره كل الترتيبات الممكنة لمجموعة من الأشياء ، على سبيل المثال ، قد نهتم بمعرفة ١٠ الترتيبات الممكنة لجلوس ستة أشخاص على مائدة

مستديرة . الترتيبات المختلفة تسمى بتباديل **Permutations** .

تعريف : التبديل هي ترتيب لكل أو جزء من فئة من الأشياء .

مثال (٣-٧) التباديل الممكنة لفئة الحروف a , b , c تكون :

abc acb bac bca cab cba

أي يوجد ستة من التباديل الممكنة . باستخدام نظرية (٣-١) يمكن الوصول إلى نفس النتيجة

بدون الحصول على قائمة بالترتيبات المختلفة . لدينا ثلاثة أماكن يمكن شغلها (ملئها) بسلاخروف

a , b , c . سوف يكون لدينا ثلاثة اختيارات للمكان الأول ، ولكن حيث أنه لا يمكن إحلال

العنصر الأول فإنه يوجد اختيارات للمكان الثاني ويتبقى اختيار واحد للمكان الأخير حال إتمام

اختيار الحرفين الأولين . وعليه فإن عدد التباديل هو $6 = 3 \cdot 2 \cdot 1 = 3!$. عموماً فإن n من

الأشياء المميزة يمكن ترتيبها بطرق عددها $3 \cdot 2 \cdot 1 \cdot \dots \cdot (n-1) \cdot n$ وسوف نرمز لهذا الناتج

بالرمز $n!$ الذي يسمى مضروب n (factorial n) . وعلى ذلك ثلاثة عناصر مختلفة

يمكن ترتيبها بطرق عددها $6 = 3 \cdot 2 \cdot 1 = 3!$. ويجب أن نعرف أن $1! = 1$ ، $0! = 1$.

نظرية (٣-٢) عدد تباديل n من الأشياء المميزة مأخوذة جميعاً في نفس الوقت هو $n!$.

في بعض الأحيان قد يكون الاهتمام بعدد التباديل لأشياء مميزة عددها n مأخوذة r في

كل مرة . فعلى سبيل المثال عدد تباديل الحروف a, b, c مأخوذة اثنين في كل مرة هو :

ab ba ac ca bc cb .

باستخدام نظرية (٣-١) مرة أخرى يكون لدينا مكانين يمكن شغلها بثلاثة اختيارات

للمكان الأول واختيارين للمكان الثاني . وعلى ذلك يكون عدد الطرق لشغل المكانين

هو $6 = 3 \cdot 2$. سوف نرمز لعدد التباديل لأشياء عددها r من بين أشياء عددها n بالرمز $P(n, r)$.

نظرية (٣-٣) عدد تباديل n من الأشياء المميزة مأخوذة r في كل مرة هو :-

$$P(n, r) = n \cdot (n-1) \cdot \dots \cdot (n-r+1) = \frac{n!}{(n-r)!}.$$

يراد أحيانا معرفة عدد تباديل مجموعة من الأشياء يكون بعضها متماثلا وتنتج الصيغة العامة لهذا العدد من النظرية التالية.

نظرية (٣-٤) عدد التباديل المختلفة لأشياء عددها n حيث n_1 من نوع و n_2 من نوع ثاني و... و n_k من النوع رقم k هو :-

$$\frac{n!}{n_1! n_2! \dots n_k!}.$$

مثال (٣-٨) كم عدد التباديل المختلفة التي يمكن تكوينها من جميع حروف كل كلمة من الكلمات التالية على حدة ؟

(١) منعم (ب) محمود

$$\text{الحل. (١)} \quad \frac{4!}{2!1!1!1!} = \frac{4!}{2!} = 12$$

لأنه يوجد 4 حروف اثنين منهم هو الحرف م .

$$(ب) \quad \frac{5!}{2!1!1!1!1!} = \frac{5!}{2!} = 60$$

لأنه يوجد 5 حروف اثنين منها هو الحرف م.

مثال (٣-٩) كم عدد الطرق التي يمكن بها ترتيب 11 كتاباً على رف إذا كان 5 كتب منهم في التاريخ و 3 في الإحصاء و 3 في الرياضيات ؟

الحل . عدد الطرق هو :-

$$\frac{n!}{n_1! n_2! n_3!} = \frac{11!}{5!3!3!} = 9240.$$

عادةً يكون اهتمامنا بعدد الطرق لتجزئة فئة من الأشياء التي عددها n إلى فئات جزئية عددها r تسمى الخلايا cells . يتحقق التجزئة إذا كان التقاطع لأي زوج من الفئات الجزئية السقي عددها r هو الفئة الخالية ϕ والاتحاد بين كل الفئات الجزئية يعطي الفئة الأصلية والترتيب للعناصر داخل الخلية ليس له أهمية. لكن الفئة $\{a, b, c, d\}$ ، التجزئات الممكنة لهذه الفئة إلى خليتين بحيث تحتوي الخلية الأولى على ثلاثة عناصر والخلية الثانية تحتوي على عنصر واحد هي :

$$\{(a, b, c, d), \{(a, b, d, c), \{(b, c, d, a), \{(a, c, d, b)\}.$$

أى أن هناك 4 طرق لتجزئة الفئة المكونة من 4 عناصر إلى خليتين تحتوي على 3 عناصر في الخلية الأولى وعنصر واحد في الخلية الثانية. عدد التقسيمات للمثال السابق يمكن كتابتها على الشكل

$$\frac{4!}{3!1!} = 4.$$

نظرية (٣-٥) عدد الطرق لتجزئة فئة من n من الأشياء إلى r من الخلايا بعناصر عددها n_1 في الخلية الأولى و n_2 من العناصر في الخلية الثانية و ... و n_r من العناصر في الخلية رقم r يكون :-

$$\frac{n!}{n_1! n_2! \dots n_r!}.$$

حيث :-

$$n_1 + n_2 + \dots + n_r = n$$

مثال (٣-١٠) بكم طريقة يمكن توزيع 10 كتب على 3 تلاميذ بحيث يتلقى التلميذ المتفوق 4 كتب وكل تلميذ آخر 3 كتب ؟

الحل . في هذا المثال يراد معرفة عدد التجزئيات لـ 10 كتب على ثلاث خلايا تحتوي على 4,3,3 من الكتب على التوالي . من النظرية السابقة عدد التجزئيات هو :-

$$\frac{10!}{4!3!3!} = 4200.$$

مثال (٣-١١) بكم طريقة يمكن توزيع 9 أشخاص على 3 غرف في فندق حيث أن غرفتين من ذات سريرين وغرفة ذات 5 أسرة .

الحل . عدد الطرق هو :-

$$\frac{n!}{n_1! n_2! n_3!} = \frac{9!}{2!2!5!} = 756.$$

في كثير من المشاكل يكون اهتمامنا بعدد الطرق لاختيار أشياء عددها r من بين أشياء مميزة عددها n ودون اعتبار لطريقة الترتيب. هذه الاختيارات تسمى التوافيق combinations . في الحقيقة التوفيق combination هو تجزئة بخليتين، خلية تحتوي على r من الأشياء والخلية الأخرى تحتوي على $(n-r)$ من الأشياء الباقية وعدد هذه التوافيق يرمز له بالرمز $\binom{n}{r}$.

نظرية (٣-٦) عدد التوافيق لأشياء مميزة عددها n مأخوذة r كل مرة هو :

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

مثال (٣-١٢) كم عدد الطرق لاختيار 8 أشخاص لفريق كرة السلة من بين 11 ولدا ؟
الحل . عدد الطرق تكون :-

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{11!}{8!3!} = 165.$$

مثال (٣-١٣) كم عدد الطرق لاختيار عملتين من كيس يحوى على دينار و ريال و درهم و
ين و فرنك ؟

الحل . عدد الطرق هي :-

$$\binom{n}{r} = \binom{5}{2} = \frac{5!}{2!3!} = 10.$$

مثال (٣-١٤) بكم طريقة يمكن اختيار بعثة علمية تتكون من 3 رجال وسيدتين من بين 7
رجال و 5 سيدات .

الحل . يمكن اختيار 3 رجال من بين 7 رجال بطرق عددها $\binom{7}{3} = 35$ ويمكن اختيار
السيدتين من بين 5 سيدات بطرق عددها $\binom{5}{2} = 10$ وبذلك يكون اختيار البعثة بطرق
عددها :-

$$\binom{7}{3} \binom{5}{2} = 350.$$

Probability

(٣-٣) الاحتمال

عادة، يكون هدف الإحصائي الوصول إلى استنتاجات أو استدلالات من التجارب التي
تشتمل على أحداث غير مؤكدة. ولكي تكون استنتاجاته واستدلالاته دقيقة فإن فهمه لنظرية
الاحتمالات يكون ضروري. ماذا نعني عندما نقول مثلا أنه من المحتمل أن تمطر السماء بعد ظهر
اليوم أو أن الاحتمال ضئيل في أن ينجح طالب معين في الامتحان... الخ. في كل حالة نعر عن
نتيجة غير متأكدين منها ولكن من معلوماتنا السابقة أو فهمنا لطبيعة التجربة يكون لدينا درجة
من الثقة في صحة هذه الجملة .

نمذنا نظرية الاحتمالات بفتة من الأرقام تسمى الأوزان **weights** تتراوح من الصفر إلى
الواحد الصحيح والتي نمكنا من تقدير لإمكانية (فرصة) وقوع الأحداث التي تنتج من تجارب
إحصائية. لكل نقطة في فضاء العينة نعين وزن بحيث يكون مجموع الأوزان يساوى الواحد
الصحيح . إذا كان لدينا سبب لكي نعتقد أن هناك إمكانية كبيرة لوقوع نقطة في فضاء العينة

فإننا نعين لها رقماً قريباً من الواحد الصحيح. ومن ناحية أخرى يعين وزن قريب من الصفر لنقاط العينة التي إمكانية وقوعها ضئيل. للنقاط خارج نطاق العينة، أى النقاط التي يستحيل حدوثها فإننا نعين لها الرقم صفر وتسمى الأحداث المستحيلة الحدوث. على سبيل المثال احتمال ظهور الرقم 7 عند إلقاء زهرة تزد مرة واحدة يساوى صفر. لإيجاد احتمال أى حدث A فإننا نجمع كل الأوزان المعنية لنقاط العينة في A . هذا المجموع يسمى المقياس $measure$ للحادثة A أو احتمال A ويرمز له بالرمز $P(A)$. الطرق المختلفة لقياس الاحتمالات تمثل مفاهيم مختلفة. في هذا البند سوف نناقش ثلاثة مفاهيم مختلفة لقياس الاحتمالات وهي: المفهوم القديم (المفهوم الكلاسيكي $classical\ concept$) و مفهوم التكرار النسبي $relative\ frequency\ concept$ والمفهوم الشخصي $subject\ concept$.

(٣-٣-١) المفهوم القديم (المفهوم الكلاسيكي) $Classical\ Concept$

تعا لهذا المفهوم تحدد أرقام الاحتمالات أو يمكن تقديرها قبلي $a\ priori$ (قبل الحقيقة $before\ fact$) ، وعلى ذلك، الاحتمال بالضبط $exact\ probability$ أن حادثة ما تقع تحدد قبل وقوع الحادثة. لذلك يسمى الاحتمال المقدر بناء على هذا المفهوم بالاحتمال القبلي $a\ priori\ probability$. المفهوم القديم للاحتمال مبنى على أساس أنه إذا كانت تجربة تحتوى على M من النقاط، أى أن عدد النواتج الممكنة لتجربة ما هو M وكانت هذه النواتج متساوية في إمكانية الحدوث وإذا احتوت الحادثة A على عدد m من النقاط فإن احتمال الحادثة هو :

$$P(A) = \frac{m}{M}$$

مثال (٣-١٥) أوجد احتمال ظهور عددين مجموعهما ثمانية عند إلقاء زهرتي لورد مرة واحدة.

الحل . فراغ العينة هو فئة الأزواج المرتبة الآتية :-

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

وكل زوج مرتب من هذه الفئة يمثل أحد نتائج التجربة. فمثلا العنصر (1,6) يمثل ظهور العدد 1 على الزهرة الأولى والعدد 6 على الزهرة الثانية. فراغ العينة يحتوى على ستة وثلاثين زوجاً مرتباً. المجموع ثمانية على الترتين سينتج إذا ظهر أى زوج من الأزواج الخمسة التالية :-

(2,6) , (6,2) , (4,4) , (5,3) , (3,5)

أى أن "ظهور عددين مجموعهم يساوى ثمانية" ينتج من 5 نقاط عينة ، وعلى ذلك :-

$$P(A) = \frac{m}{M} = \frac{5}{36}$$

مثال (٣-١٦) أسرة لديها 3 أطفال. ما احتمال أن يكون جميعهم أولاد (بفرض أن كل طفل له نفس الاحتمال أن يكون ولدا أو بنتا).

الحل . إذا رمزنا للولد بالرمز B وللبنات بالرمز G فإن فراغ العينة هو :-

$$S = \{BBB, BBG, BGB, GBB, GBG, BGG, GGB, GGG\}$$

الحادثة " جميع الأطفال أولاد " هو $A = \{BBB\}$ ، أى أن $M = 8$, $m = 1$ وبالتالي فإن:-

$$P(A) = \frac{m}{M} = \frac{1}{8}$$

Relative Frequency Concept

(٣-٣-٢) مفهوم التكرار النسبي

يشترط هذا المفهوم إجراء التجربة عدد كبير من المرات ومعرفة نتائجها وبعد ذلك قياس الاحتمال. فإذا كانت N تمثل عدد المرات (المحاولات $trails$) التي أجريت بها تجربة ما تحت نفس الظروف و n تمثل عدد مرات (التكرار) ظهور الحادثة A خلال N من المرات السبي كورت فيها التجربة فإن احتمال وقوع الحادثة A هو :-

$$P(A) = \lim_{N \rightarrow \infty} \frac{n}{N}$$

حيث $\frac{n}{N}$ هو التكرار النسبي للحادثة A في هذه التجارب التي عددها N . عادة تكون قيم

التكرار النسبي غريبة الأطوار للقيم الصغيرة من N ولكن عندما تزيد قيمة N ، فقد أوضحت

الخبرة أن النسبة $\frac{n}{N}$ تكتسب بعض الانتظام الإحصائي وتستقر حول قيمة ثابتة هي $P(A)$

ولذلك عرف الاحتمال بأنه نهاية النسبة $\frac{n}{N}$ عندما يزداد عدد المحاولات أو التجارب ويؤول إلى

ما لا نهاية. وحيث أن عدد المحاولات نادرا ما يؤول إلى ما لا نهاية فإننا نستخدم التكرار النسبي

كتقدير لقيمة الاحتمال. الاحتمال المبني على هذا المفهوم يقدر بعدى $a posteriori$ (بعد

الحقيقة $after fact$) ويعتمد على البيانات الملاحظة، لذلك يسمى في بعض الأحيان الاحتمال

التجريبي $experiment probability$.

مثال (٣-١٧) في مصنع لإطارات السيارات تبين أن كل 100000 إطار منتج يكون من بينها

300 إطار تالف. فما احتمال اختيار إطار تالف؟

الحل . عدد الإطارات $N=100000$ عدد الإطارات التالفة $n = 300$. وعلى ذلك احتمال اختيار إطار تالف هو :-

$$P(A) \approx \frac{300}{100000} = 0.003.$$

Subject Probability

(٣-٣-٣) المفهوم الشخصي

تعبّر هذا المفهوم، الاحتمال هو درجة الثقة **degree of confidence** في وقوع حادثة ما والمقررة من شخص ما بناء على دليل متوفر لديه **evidence available** . هذا الدليل قد يكون أي معلومات كمية أو غير كمية. على سبيل المثال قد يحدد الشخص القسام على المشتريات في شركة ما الاحتمال 0.25 للحادثة أن شحنة ما تحتوي على أكثر من 2% وحدات تالفة . أيضا قد يصرح المدير الأول في شركة بأن احتمالات زيادة ميزانية الشركة أو انخفاضها أو بقائها على حالها هو 0.04 و 0.13 و 0.83 على التوالي. ويجب ملاحظة أن الاحتمال المقدر لحادثة ما بناء على هذا المفهوم يختلف من شخص إلى آخر وذلك لعوامل كثيرة منها الخبرة .

(٤-٣-٣) الخواص المميزة لقيم الاحتمال

Characteristics of Probability Numbers

إذا كان S فراغ العينة المرافق لتجربة وإذا كانت A_1, A_2, \dots تمثل كسل الأحداث الممكنة فإن قيم الاحتمال المقدرة للأحداث السابقة لابد أن تتوافر فيها الشروط الآتية :-

(أ) يرافق كل حادثة A عدد معين $P(A)$ يسمى احتمال A ويحقق $P(A) \geq 0$.

(ب) احتمال وقوع حادثة مؤكدة يساوي الواحد الصحيح، أي أن $P(S) = 1$

(جـ) إذا كانت A_1, A_2, A_3, \dots عدد لإنهائي من الحوادث المانعة بالتبادل أي $A_i \cap A_j = \emptyset$, $i \neq j$ فإن :-

$$P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$$

ويمكن إثبات أنه إذا كانت A_1, A_2, \dots, A_n تمثل حادثة مانعة بالتبادل فإن :-

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

أيضا إذا كانت A_1, A_2, \dots, A_n تمثل تجزئة لفراغ العينة S فإن :

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) = 1.$$

مثال (١٨-٣) ما هو احتمال الحصول على مجموع 5 أو 12 عند إلقاء زهرتين نرد مرة واحدة.

الحل . بفرض أن A حادثة "ظهور 5" و B حادثة "ظهور 12" . المجموع 5 يحدث من 4 نقط عينة والمجموع 12 يحدث من نقطة واحدة. وحيث أن كل النتائج في فضاء العينة متساوية في إمكانية الحدوث فإن $P(A) = \frac{4}{36}$ ، $P(B) = \frac{1}{36}$ ، الحادتان A ، B متالعتان وعلى ذلك :-

$$P(A \cup B) = P(A) + P(B) \\ = \frac{4}{36} + \frac{1}{36} = \frac{5}{36}.$$

(٣-٤) بعض قوانين الاحتمال Some Probability Laws

عادة يكون من السهل حساب احتمال حادثة ما من الاحتمالات المعروفة للأحداث الأخرى وهذا يكون صحيح إذا أمكن تمثيل الحادثة كاتحاد لحادتين أخريتين أو مكملته لحادثة. في هذا البند سوف نعرض بعض القوانين التي تسهل عملية حساب الاحتمالات. نظرية (٣-٧) لأى حادتين A ، B فإن :-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

البرهان :-

كما يتضح من شكل (٣-١) أن $I = A^c \cap B$ ، $II = A \cap B^c$ ، $III = A \cap B$ ، يمكن تمثيل كل من الحادتين $A \cup B$ و A كاتحاد لحادتين متالعتان كما يلي :-

$$A \cup B = (A \cap B^c) \cup B , \\ A = (A \cap B) \cup (A \cap B^c).$$

حيث :

$$(A \cap B^c) \cap B = \phi , \\ (A \cap B) \cap (A \cap B^c) = \phi.$$

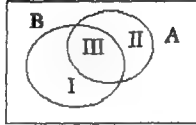
وعلى ذلك :-

$$P(A \cup B) = P(A \cap B^c) + P(B) , \\ P(A) = P(A \cap B) + P(A \cap B^c).$$

وبذلك نصل إلى :-

$$P(A \cup B) = P(A \cap B^c) + P(B) \\ = [P(A) - P(A \cap B)] + P(B)$$

$$= P(A) + P(B) - P(A \cap B).$$



شكل (٣-١)

مثال (٣-١٩) وجد أن 40% من المرضى الذين تم فحصهم في العيادة انجليزية يعانون من ارتفاع في ضغط الدم ، وأن 30% يعانون من زيادة الوزن وأن 10% يعانون من كليهما. إذا اختبر مريضاً عشوائياً ، فما هو احتمال أن يعاني من إحدى هاتين الحالتين على الأقل ؟

الحل . بفرض A حادثة " ارتفاع ضغط الدم " و B حادثة " الوزن زائد " و $A \cup B$

حادثة " المعاناة من إحدى هاتين الحالتين " . من نظرية (٣-٧) :-

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \\ = 0.4 + 0.3 - 0.1 = 0.6.$$

نظرية (٣-٨) إذا كانت A حادثة وكانت A^c الحادثة المكملية لها فإن :-

$$P(A) = 1 - P(A^c).$$

البرهان :-

نعلم أن A^c هي المكملية للفترة A وعلى ذلك $S = A \cup A^c$. وحيث أن :

$$A \cap A^c = \phi \text{ لأن } A, A^c \text{ حادتين متامعتان ، فإن :-}$$

$$P(S) = P(A \cup A^c)$$

إذن :

$$P(S) = P(A) + P(A^c)$$

$$1 = P(A) + P(A^c).$$

وعلى ذلك :-

$$P(A) = 1 - P(A^c).$$

مثال (٣-٢٠) إذا أُلقيت عملة 7 مرات أوجد احتمال ظهور صورة مرة على الأقل.

الحل . بفرض A الحادثة " ظهور صورة مرة على الأقل " . عدد نقط العينة في S هو 2^7 (من نظرية (٣-١)) . بفرض أن الحادثة A^c تمثل عدم ظهور أي صورة وحيث أن الأحداث متساوية في إمكانية الحدوث فإن :

$$P(A^c) = \frac{1}{2^7} = 0.0078125.$$

أي أن :-

$$\begin{aligned} P(A) &= 1 - P(A^c) \\ &= 1 - 0.0078125 = 0.9921875. \end{aligned}$$

(٣-٥) الاحتمال الشرطي Conditional Probability

في بعض التجارب يتأثر الاحتمال الذي يخص حادثة ما (لنكن A) بالمعلومات عن حدوث أو عدم حدوث حادثة أخرى ولنكن B . في هذه الحالة سوف نستخدم العبارة : احتمال وقوع حادثة A بشرط وقوع حادثة B والذي يسمى الاحتمال الشرطي ويرمز له بالرمز $P(A|B)$ ويقرأ " احتمال وقوع A الشرط وقوع B " . للتسهيل بفرض أن A حادثة الحصول على رقم 3 عند إلقاء زهرة متزنة مرة واحدة. بالاعتماد على فضاء العينة $S = \{1,2,3,4,5,6\}$ فإن احتمال الحصول على الرقم 3 عند إلقاء زهرة نرصد مرة واحدة هو $\frac{1}{6}$. الآن بفرض أن الشخص الذي قام بإلقاء زهرة الترد ألدانا بأن النتيجة التي حصل عليها كانت رقم فردي . لنعرض آثار هذه المعلومات التي توفرت مسبقاً على احتمال الحادثة A والتي حسبناها مسبقاً. الآن في ظل هذه الحقيقة لم تعد النتائج الممكنة ستة وإنما أصبحت ثلاثة فقط فهي إما 1 أو 3 أو 5. أما النتائج 2 أو 4 أو 6 فأصبحت مستحيلة. وعلى ذلك فإن احتمال الحادثة A منسوباً إلى الفراغ الجديد المختزل $B = \{1,3,5\}$ هو :

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{1}{3}.$$

حيث $n(A \cap B)$, $n(B)$ تمثل عدد العناصر التابعة للحادثتين $A \cap B$, B على التوالي . مثال (٣-٢١) ألقيت عملة ثلاث مرات فإذا علم أن الوجه الظاهر في الرمية الأولى والثانية كتابة ما هو احتمال ظهور كتابة في الرمية الثالثة ؟

الحل . فراغ العينة هو :-

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

بفرض أن الحادثة $A = \{HHT, THT, HTT, TTT\}$ "ظهور كتابة في الرمية الثالثة" و
 الحادثة $B = \{TTH, TTT\}$ "ظهور كتابة في الرمية الأولى والثانية".
 الحادثة $A \cap B = \{TTT\}$ تحتوي على نقطة واحدة والحادثة B تحتوي على نقطتين.
 باستخدام فراغ العينة المختزل B ، وإذا كانت $n(A \cap B)$ تمثل عدد العناصر التابعة
 للحادتين $A \cap B$ ، B على التوالي. وعلى ذلك يكون الاحتمال المطلوب هو :

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{1}{2}.$$

مثال (٣-٢٢) وعاء يحتوي على 150 وحدة بعضها منتج من المصنع 1 والباقي من المصنع 2، بعض الوحدات سليمة وبعضها تالفة. فإذا اختيرت وحدة عشوائية من الوعاء،
 ليكن A الحادثة "الوحدة تالفة" وعلى ذلك، A^c الحادثة "الوحدة سليمة". أيضا ليكن B
 الحادثة "الوحدة من إنتاج المصنع 1"، B^c الحادثة "الوحدة من إنتاج المصنع 2". الجدول
 التالي يمثل عدد الوحدات السليمة والتالفة المنتجة من المصنعين. الآن بفرض أن كل وحدة
 مرقمة بعلامة توضح المصنع المنتج أوجد $P(A|B)$.

	B	B^c	المجموع
A	50	5	55
A^c	75	20	95
المجموع	125	25	150

الحل. يمكن حساب احتمال أن الوحدة تالفة بشرط وقوع الحادثة B وباستخدام القسراغ
 المختزل B كالآتي :-

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{50}{125} = \frac{2}{5}.$$

أيضا يمكن كتابة $P(A|B)$ على الشكل :

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{n(A \cap B) \cdot n(S)}{n(B) \cdot n(S)} = \frac{P(A \cap B)}{P(B)}.$$

حيث $P(B)$, $P(A \cap B)$ يتم الحصول عليهما من فراغ العينة الأصلي S . للتحقق من النتيجة فإن :-

$$P(B) = \frac{125}{150} ,$$

$$P(A \cap B) = \frac{50}{150} .$$

وعلى ذلك :-

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{50}{125} = \frac{2}{5} .$$

تعريف : الاحتمال الشرطي للحادثة A شرط B يمثل بالصيغة $P(A|B)$ و يعرف بالمعادلة :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} , P(B) \neq 0 .$$

يعتبر هذا التعريف عام ولا يعتمد على فراغ عينة يحتوى على أحداث متساوية في إمكانية الحدوث كما في الأمثلة السابقة بنفس الشكل يمكن القول أن :

$$P(B|A) = \frac{P(A \cap B)}{P(A)} , P(A) \neq 0 .$$

مثال (٣ - ٣٣) في إحدى الكليات نجح % 75 من الطلبة في امتحان الرياضيات ونجح % 85 من الطلبة في امتحان الإحصاء ونجح % 10 في الرياضيات والإحصاء . اختصر أحد الطلبة بطريقة عشوائية . المطلوب :-

- (أ) إذا كان ناجحاً في الإحصاء ما هو احتمال أن يكون ناجحاً في الرياضيات .
 (ب) إذا كان ناجحاً في الرياضيات فما هو احتمال أن يكون ناجحاً في الإحصاء .
 الحل . بفرض أن A حادثة " النجاح في الرياضيات " و B حادثة " النجاح في الإحصاء " وعلى ذلك :-

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.1}{0.85} = \frac{10}{85} = \frac{2}{17} \quad (أ)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.1}{0.75} = \frac{10}{75} = \frac{2}{15} \quad (ب)$$

نظرية (٣ - ٩) إذا وقعت حادثة ما A في تجربة ما، يتبعها حادثة B فإن :-

$$P(A \cap B) = P(A)P(B|A) .$$

وعلى ذلك احتمال وقوع B , A في ترتيب هو احتمال أن تقع A أولاً مضروباً في احتمال وقوع B , شرط أن A وقعت . كما يمكن أن يكون :-

$$P(A \cap B) = P(B)P(A|B).$$

وهذا يتوقف على أى الحادتين قد تقع أولاً .

مثال (٣-٢٤) كيس يحتوي على 4 كرات بيضاء و 7 حمراء فإذا أختار شخص كرتين من الكيس اختياراً عشوائياً فما احتمال أن تكون الأولى بيضاء والثانية حمراء ؟ (إذا كان السحب بدون إرجاع) .

الحل . بفرض أن A الحادثة "الكرة الأولى بيضاء" و B الحادثة "الكرة الثانية حمراء" وعلى ذلك احتمال أن تكون الكرة الأولى بيضاء والثانية حمراء هو :-

$$P(A \cap B) = P(A)P(B|A) = \left(\frac{4}{11}\right) \left(\frac{7}{10}\right) = \frac{14}{55}.$$

نظرية (٣-١٠) في أى تجربة إذا وقعت الحادثة A_1 ، يتبعها الحادثة A_2 ، يتبعها الحادثة A_3 ، وهكذا ، فإن :-

$$P(A_1 \cap A_2 \cap A_3 \cap \dots) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \dots$$

في بعض الأحيان ، احتمال وقوع حادثة ما A لا يتأثر ولا يعتمد على وقوع أو عدم وقوع حادثة أخرى B . بعبارة أخرى وفي هذه الحالة يقال أن A مستقلة عن B وعلى ذلك $P(A|B) = P(A)$ وفي هذه الحالة فإن :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A).$$

ومنها :

$$P(A \cap B) = P(A)P(B).$$

إذا كانت A مستقلة عن B فإن B تكون مستقلة عن A لأن :-

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B).$$

ومنها :

$$P(A \cap B) = P(A)P(B).$$

تعريف : يقال أن الحادتين A ، B مستقلتين **independent** ، إذا وفقط إذا :-

$$P(A \cap B) = P(A)P(B).$$

مثال (٣-٢٥) في المثال السابق إذا كان السحب بإرجاع المطلوب إيجاد احتمال أن تكون الكرة الأولى بيضاء والثانية حمراء؟

الحل . بفرض أن A الحادثة " الكرة الأولى بيضاء " ، B الحادثة " الكرة الثانية حمراء " الحادتين A و B مستقلتين وعلى ذلك احتمال أن تكون الكرة الأولى بيضاء والثانية حمراء هو:-

$$P(A \cap B) = P(A)P(B) = \left(\frac{4}{11}\right) \left(\frac{7}{11}\right) = \frac{28}{121}$$

مثال (٢٦-٣) ألقيت عملة وزهرة نرد معا، ما هو احتمال ظهور الصورة على العملة والرقم 6 على زهرة النرد ؟

الحل . بفرض أن الحادثة A "ظهور الصورة على العملة " والحادثة B ظهور رقم 6 على النرد . ولأن الحادتين مستقلتين فإن :-

$$P(A \cap B) = P(A)P(B) \\ = \left(\frac{1}{2}\right) \left(\frac{1}{6}\right) = \frac{1}{12}$$

مثال (٢٧-٣) ألقيت زهرتي نرد مرة واحدة ما هو احتمال ظهور الرقم 5 على الزهرة الأولى والرقم 2 على الزهرة الثانية ؟

الحل . بفرض أن A الحادثة ظهور الرقم 5 على الزهرة الأولى، B الحادثة ظهور الرقم 2 على الزهرة الثانية . بما أن الحادتين مستقلتين فإن:-

$$P(A \cap B) = P(A)P(B) \\ = \left(\frac{1}{6}\right) \left(\frac{1}{6}\right) = \frac{1}{36}$$

مثال (٢٨-٣) ألقيت زهرتي نرد مرتين . ما هو احتمال أن مجموع الوجهين 5 في رمية و7 في الرمية الأخرى ؟

الحل . بفرض أن A_1, A_2, B_1, B_2 على التوالي ظهور مجموع 5 في الرمية الأولى و مجموع 5 في الرمية الثانية و ظهور مجموع 7 في الرمية الأولى و مجموع 7 في الرمية الثانية (تمثل أحداث مستقلة) . اهتمامنا سوف يكون في حساب احتمال الاتحاد لحادتين مانعتين

:- وعلى ذلك ، $B_1 \cap A_2, A_1 \cap B_2$

$$\begin{aligned} P[(A_1 \cap B_2) \cup (B_1 \cap A_2)] &= P(A_1 \cap B_2) + P(B_1 \cap A_2) \\ &= P(A_1)P(B_2) + P(B_1)P(A_2) \\ &= \left(\frac{4}{36}\right) \left(\frac{6}{36}\right) + \left(\frac{6}{36}\right) \left(\frac{4}{36}\right) = 0.037037. \end{aligned}$$

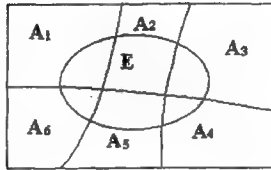
(٢٩-٣) الاحتمال الكلي وقاعدة بييز

بفرض أن الأحداث A_1, A_2, \dots, A_n تمثل تجزيتنا لفراغ العينة وماعة لبعضها البعض

والتحادهم هو S (أحداث مائعة وشاملة mutually exclusive and exhaustive)

كما في شكل (٣-٢) حيث $n=6$ • بفرض أن أي حادثة أخرى فإن :-

$$\begin{aligned} E &= S \cap E = (A_1 \cup A_2 \cup \dots \cup A_n) \cap E \\ &= (A_1 \cap E) \cup (A_2 \cap E) \cup \dots \cup (A_n \cap E) \end{aligned}$$



شكل (٣-٢)

الصيغة السابقة مفيدة في النظرية الآتية :-

نظرية (٣-١١) (نظرية الاحتمال الكلي total probability)

بفرض أن A_1, A_2, \dots, A_n تمثل n حادثة مائعة وشاملة، وعلى ذلك لأي حادثة E

فإن :-

$$P(E) = \sum_{i=1}^n P(A_i)P(E|A_i).$$

البرهان :-

الأحداث $A_1 \cap E, A_2 \cap E, \dots, A_n \cap E$ مائعة بالتبادل، وعلى ذلك :-

$$P(E) = \sum_{i=1}^n P(A_i \cap E),$$

ويتطبيق نظرية (٣-٩) على كل حد نحصل على برهان النظرية •

مثال (٣-٢٩) تنتج ثلاث ماكينات A, B, C ، 40% ، 35% ، 25% على التوالي

من الإنتاج الكلي لمصنع، ونسبة الإنتاج السليم لهذه الماكينات هي 98% ، 96% ، 94%

فإذا اختبرت وحدة بطريقة عشوائية، ما هو الاحتمال أن تكون سليمة ؟

الحل . بفرض أن الحادثة E "الوحدة المختارة سليمة" ، والحادثة A_1 "الوحدة من إنتاج الماكينة A" والحادثة A_2 "الوحدة من إنتاج الماكينة B" والحادثة A_3 "الوحدة من إنتاج الماكينة C" وعلى ذلك :-

$$P(E) = \sum_{i=1}^3 P(A_i)P(E|A_i)$$

$$= (0.4)(0.98) + (0.35)(0.96) + (0.25)(0.94) = 0.9630 .$$

نظرية (٣-١٢) نظرية Bayes` Theorem

إذا كانت A_1, A_2, \dots, A_n تمثل n حادثة مانعة وشاملة وكان ظهور إحداها ينتج عنه ظهور حادثة أخرى E (أى أن E تقع إذا وقعت واحدة من الحوادث المانعة) فإن :-

$$P(A_k|E) = \frac{P(A_k)P(E|A_k)}{\sum_{i=1}^n P(A_i)P(E|A_i)}, k=1,2,\dots,n.$$

البرهان :-

نعلم من النظرية السابقة أن :-

$$P(E) = \sum_{i=1}^n P(A_i)P(E|A_i).$$

وحيث أن :-

$$P(A_k|E) = \frac{P(A_k \cap E)}{P(E)} = \frac{P(A_k)P(E|A_k)}{P(E)}.$$

وبالتعويض عن $P(E)$ بقيمتها من نظرية (٣-١١) يتم البرهان .

مثال (٣-٣٠) تنتج إحدى شركات المشروبات نوع معين من العصائر . يستمر الإنتاج خلال ورديتين بحيث أن 70% من الإنتاج اليومي من الوردية الأولى . من دراسة المنتج وجد أن نسبة العبوات السليمة من إنتاج الوردية الأولى 95% ونسبة العبوات السليمة من إنتاج الوردية الثانية 97% . فإذا سُحبت إحدى العبوات عشوائيا وكانت سليمة ما هو احتمال أن تكون من إنتاج الوردية الثانية ؟

الحل . بفرض أن E الحادثة "العبوة سليمة" و A_1 الحادثة "العبوة من إنتاج الوردية الأولى" و A_2 الحادثة "العبوة من إنتاج الوردية الثانية" وعلى ذلك الاحتمال المطلوب هو :-

$$P(A_2 | E) = \frac{P(A_2)P(E | A_2)}{P(A_1)P(E | A_1) + P(A_2)P(E | A_2)}$$

$$= \frac{(0.3)(0.97)}{(0.7)(0.95) + (0.3)(0.97)} = 0.3043933.$$

تمارين

١ - ألقى زوج من زهرتي الترد مرة واحدة أذكر الحادثة :-

(أ) مجموع الوجهين الظاهرين يساوي 9 .

(ب) مجموع الوجهين الظاهرين إما 4 أو 5 .

٢ - ألقيت عملتين مرة واحدة أذكر الحادثة :-

(أ) ظهور كتابة واحدة .

(ب) ظهور كتابة واحدة على الأقل .

٣ - في تجربة لاختبار ثلاث وحدات من مصنع وملاحظة ما إذا كانت الوحدة سليمة أو تالفة (

يرمز للتالفة D والسليمة D') أذكر :-

(أ) فضاء العينة .

(ب) الحادثة عدم ظهور وحدات تالفة .

(ج) كيف يمكن تعريف الحادثة ؟

$$A = \{(DDD'), (DD'D), (D'DD)\}$$

٤ - اختبر أربعة أشخاص عشوائيا لاختبار تفضيل أو عدم تفضيل لنوع معين من القهوة حيث

يعطى 1 للتفضيل و 0 لعدم التفضيل أذكر :-

(أ) فضاء العينة .

(ب) الحادثة ثلاث أشخاص على الأقل يفضلون .

٥ - قام مسئول بمراقبة الجودة في مصنع لإنتاج أسماك السلمون باختبار كل صندوق منتج وأخذ

عينة والاستمرار في الاختبار حتى ظهور صندوق تالف . أذكر فضاء العينة لعملية الاختبار مع العلم

أن Y ترمز للصندوق السليم و N ترمز للصندوق التالف .

٦ - قام باحث متخصص في التسويق بتصنيف العملاء إلى ثلاث مجموعات حسب الدخل:

منخفض 0 ومتوسط 1 وعالي 2 . كما قام بتصنيفهم تبعاً لخاصية أخرى وهي القوة الشرائية إلى

(لا يشتري 0) و (يشتري ولو مرة واحدة في الشهر 1) . عرف فضاء العينة .

٧ - بفرض عدم السماح بالتكرار (أ) كم عدداً مكون من ثلاث أرقام يمكن تركيبة من الأرقام

التالية 8,7,3,2,1 ؟ (ب) كم عدداً منهم زوجياً ؟ (جـ) كم عدداً منهم فردياً ؟

٨ - بكم طريقة يمكن أن يجلس أربعة طلاب من قسم الكيمياء وأربعة طلاب من قسم النبات وثلاثة طلاب من قسم الرياضيات وطلابتان من قسم الحيوان في صف بحيث يجلس الأشخاص ذو التخصصات الواحدة معا ؟

٩ - حل نفس التمرين السابق إذا جلس الجميع حول مائدة مستديرة .

١٠ - كم عدد الطرق لاختيار ثلاثة عملات من صندوق يحتوي على جنيه و ريال و دينار و ين و فرنك ؟

١١ - كم عدد الطرق لاختيار ثمانية أشخاص لفريق كرة القدم من 14 شخصا ؟

١٢ - صنعت زهرة نرد بحيث يكون احتمال ظهور الرقم واحد ثلاثة أضعاف أي رقم آخر، بينما كل الوجوه الأخرى لها نفس الفرصة في الظهور . ما هو احتمال ظهور الرقم اثنين عند إلقاء النرد مرة واحدة ؟ وما هو احتمال ظهور الرقم واحد ؟

١٣ - مطلوب من طالب دراسة مادة في العلوم ومادة في الرياضيات ومادة في الاجتماع . ما هو عدد الطرق لاختيار هذه المواد من بين 3 مواد في العلوم و 4 في الاجتماع و مادتين في الرياضيات ؟

١٤ - ما عدد الطرق الممكنة لشخص داخل محل ملابس لاختيار رابطة عنق و قميص وذلك إذا توافر له 4 أربطة عنق و 5 قمصان في المحل ؟

١٥ - بكم طريقة يمكن زراعة 8 شجرات على شكل دائرة ؟

١٦ - كم عددا مكون من ثلاثة أرقام التي يمكن تكوينها من الأعداد 0,1,2,3,4,5 ؟ وإذا كان كل رقم يظهر مرة واحدة (١) كم عدد الأرقام الفردية ؟ (ب) كم عدد الأرقام الزوجية ؟

١٧ - إذا لعب فريق كرة القدم ثمانية مباريات خلال الموسم . بكم طريقة يستطيع الفريق في نهاية الموسم أن يكسب 4 ويفقد 3 ويتعاد 1 ؟

١٨ - بكم طريقة يمكن الإجابة 10 على أسئلة من نوع صح وخطأ ؟

١٩ - أعطى امتحان في مادة الإحصاء لطلاب . يتكون الامتحان من 9 أسئلة منهم 6 أسئلة اختياري وثلاثة إجباري . فإذا كان المطلوب منه الإجابة على ستة أسئلة . بكم طريقة يمكن للطلاب اختيار الأسئلة التي يرغب الإجابة عليها ؟

٢٠ - بكم طريقة يمكن للمدرس أن يختار طالبا من بين سبعة طلاب ؟

٢١ - أوجد عدد الطرق التي يمكن بها تصنيف 5 كتب من الحجم الكبير و 4 من الحجم المتوسط و 3 من الحجم الصغير على إحدى الرفوف بشرط أن تكون جميع الكتب ذات الحجم الواحد مصفوفة معا .

- ٢٢ - ثلاث أجزاء من كتاب موضوعة على رف ما احتمال (١) الأجزاء في وضعها الصحيح ؟
(ب) الجزء الثاني في المكان الأول ؟

- ٢٣ - اختبرت ثلاثة كتب عشوائيا من رف يحوى على 5 كتب في التاريخ و3 كتب في العلوم وقاموس ما هو احتمال (١) القاموس هو المختار ؟ (ب) كتابين في العلوم و واحد في التاريخ ؟ المختاران ؟

- ٢٤ - في مدينة ما احتمال أن أسرة تشتري تلفزيون هو 0.8 واحتمال أن تشتري غسالة ملابس هو 0.5 واحتمال أن تشتري الاثنين معا هو 0.45 ، ما هو احتمال أن تشتري الأسرة واحد من الاثنين على الأقل ؟

- ٢٥ - احتمال أن تعطر السماء في بلد ما في 4 يوليو هو 0.1 واحتمال حدوث زلزال هو 0.5 واحتمال حدوث مطر أو زلزال هو 0.3 ما هو احتمال حدوث مطر وزلزال في نفس اليوم ؟

- ٢٦ - لاعب كرة يكسب 50% من مبارياته . ما هو احتمال أن يكسب بالضبط 3 من الأربع مباريات القادمة ؟

- ٢٧ - وعاء يحوى على 10 وحدات منهم 3 تألفين سحبت وحدتين من الوعاء الواحدة تلو الأخرى بدون إرجاع ، المطلوب تقدير (١) احتمال أن الوحدتين غير تألفين (ب) احتمال وجود وحدة تألفة .

- ٢٨ - صندوق يحوى على 5 كرات سوداء و3 كرات خضراء . سحبت ثلاث كرات الواحد تلو الأخرى بدون إرجاع ما احتمال أن كل الكرات المسحوبة من نفس اللون ؟

- ٢٩ - إذا كان A , B حادثتين بحيث أن :-

$$P(A \cap B) = \frac{1}{10}, P(A^c \cap B) = \frac{1}{5}, P(A \cap B^c) = \frac{2}{3}$$

(١) ما هو احتمال A^c , B^c ؟

(ب) هل الحادثتين مستقلتين ؟

- ٣٠ - مسحضر في أنبوبة اختبار يحوى على عشرين من حبوب لقاح الصنوبر وخمسة من حبوب لقاح البلوط . اختبرت عينة عشوائية تحوى على أربعة حبوب لقاح، ما هو احتمال أن :

(١) تحتوى العينة على أربع حبوب من الصنوبر .

(ب) تحتوى العينة على ثلاث حبوب من البلوط .

(جـ) تحتوى العينة على الأقل على ثلاث حبوب لقاح الصنوبر .

- ٣١ - أطلق صياد 7 طلقات نارية على حيوان مفترس فإذا كان احتمال أن يصيب الهدف هو 0.6 ما هو احتمال أن الصياد ما زال على قيد الحياة ؟

٣٢ - صوب شخصان ناحية هدف ما، فإذا كان في المتوسط A يكسب 3 من 5 و B يكسب 4 من 8 ما هو احتمال أن الهدف لا يستهدف إذا صوب الاثنين ناحية الهدف ؟

٣٣ - تقدم ثلاثة أشخاص A , B , C إلى ثلاث وظائف مختلفة . احتمال أن يكسب A الوظيفة هو 0.8 واحتمال أن يكسب B الوظيفة هو 0.5 واحتمال أن يكسب C الوظيفة 0.45 . ما هو احتمال (١) أن الثلاثة يحصلون على الوظائف ؟ (ب) عدم تعيين أى واحد في الوظيفة المقدم لها ؟ (جـ) واحد فقط يحصل على الوظيفة ؟

٣٤ - مجموعة مكونة من عشرة أشخاص، منهم ستة أشخاص مصابين بـسوس الأسنان . اختبرت منهم عينة عشوائية من ثلاثة أشخاص . ما هو احتمال أن تحتوى العينة على ثلاثة أشخاص مصابين بـسوس الأسنان .

٣٥ - آلة تتكون من ثلاثة أجزاء، الآلة تعبر تالفة إذا كان واحد أو أكثر تالفاً . احتمال أن الجزء A يتلف هو 0.01 واحتمال الجزء B يتلف هو 0.02 واحتمال الجزء C يتلف هو 0.1 أوجد احتمال (١) الآلة تالفة (ب) أن تلف الآلة يرجع إلى فشل الجزء C فقط .

٣٦ - شركة طيران لها ست رحلات من بلد A إلى B وسبع رحلات من B إلى C (يومياً) ما عدد الرحلات التي تنجزها يومياً من A إلى C ؟

٣٧ - اختبرت ثلاثة فئران من مجموعة مكونة من خمسة فئران بيضاء اللون وأربعة بيضاء اللون لاستخدامها في تجربة معينة . ما هو احتمال أن تكون :-

(١) جميع الفئران المختارة بيضاء اللون ؟

(ب) الفئران المختارة مكونة من فأر بني وفأرين لونهما أبيض ؟

(جـ) جميع الفئران المختارة بيضاء اللون ؟

٣٨ - اختبرت ثلاث بذرات نبات مزهر عشوائية من كيس يحتوى على عشرة بذور زهورها حمراء وخمس زهورها بيضاء، ما هو احتمال أن تكون :-

(١) زهور البذور الثلاثة من نفس اللون ؟

(ب) زهور البذور الثلاثة المختارة مختلفة الألوان ؟

٣٩ - في مستعمرة كبيرة لذباب الفاكهة ، 20% من الذباب به طفرة في الجناح، 35% بسه طفرة في العين، 10% به طفرة بكل من الجناح والعين . اختبرت ذبابة من المستعمرة عشوائية . ما هو احتمال أن يكون لها أحد الطفرتين على الأقل ؟

٤٠ - حققت ثمانية فئران بمقار معين وتم رصد عدد الفئران التي ماتت خلال يسوم . إذا كان احتمال موت ستة بالضبط هو 0.03 واحتمال أن يموت سبعة أو ثمانية هو 0.04 . أوجد احتمال أن :-

(أ) يموت ستة لفران أو أكثر .

(ب) يموت خمسة أو أقل .

- ٤١ - إذا عُلِمَ أن احتمال أن يكون الجو في بلد ما في شهر يناير ملبداً بالغيوم هو 0.6 واحتمال أن يكون الجو عاصفاً هو 0.65 واحتمال أن يكون ملبداً بالغيوم وعاصفاً هو 0.25 أوجد الاحتمالات الآتية :-

(أ) أن يكون الجو ملبداً بالغيوم وغير عاصف .

(ب) أن يكون غير ملبد بالغيوم وغير عاصف .

(جـ) أن يكون الجو غير ملبد بالغيوم وعاصف .

- ٤٢ - بفرض أن 100 مستودع تم تصنيفهم حسب الإدارة (A,B,C) وحسب المبيعات (عالي، متوسط، منخفض) في الجدول المزدوج التالي :-

الإدارة				
المبيعات	A	B	C	المجموع
عالي	20	4	2	26
متوسط	4	46	14	64
منخفض	1	2	7	10
المجموع	25	52	23	100

أستخدم البيانات الموجودة في الجدول في حساب :-

(أ) احتمال أن المبيعات عالية .

(ب) احتمال أن المبيعات متوسطة إذا علم أن الإدارة A .

(جـ) احتمال أن المبيعات متوسطة إذا علم أن الإدارة B .

- ٤٣ - في استطلاع للرأي عن تأثير الإعلانات على البيع في مركز لتسويق الأغذية، أخذت عينة من 230 فرد من المترددين على المركز وسُجِلَت إجاباتهم . الجدول المزدوج التالي يوضح توزيع الأفراد حسب الشراء (يشتري ولا يشتري) وحسب مشاهدة الإعلانات (يشاهد ولا يشاهد) .

سحبت استمارة عشوائياً فإذا علم أن الشخص المختار يشاهد الإعلانات فما هو احتمال أن يشتري .

المجموع	لا يشترون	يشترون
180	100	80
50	30	20
230	130	100

- ٤٤ - في بحث ميداني لدراسة العلاقة بين العمر و استخدام حزام الأمان في مدينة بما 1000
وقائد سيارة تم الحصول على البيانات التالية :-

المجموع	يستخدمون الحزام	لا يستخدمون الحزام	
427	177	250	تحت 40
573	248	325	40 فأكثر
1000	425	575	المجموع

(أ) ما هو احتمال أن قائد السيارة يستخدم حزام الأمان ؟

(ب) ما هو احتمال أن قائد السيارة تحت 40 سنة ولا يستخدم حزام الأمان ؟

(جـ) إذا علم أن قائد السيارة يستخدم حزام الأمان ما هو احتمال أن عمره 40 فأكثر .

- ٤٥ - في متجر لبيع الملابس النسائية تم تصنيف 230 فرد من المترددين على المتجر حسب

الشراء إلى (يشترون ولا يشترون) وحسب الجنس (ذكور إناث) كما في الجدول المزدوج

التالي :-

(أ) ما هو الاحتمال أن المشتري أنثى ؟

(ب) إذا تم الشراء ما هو احتمال أن المشتري أنثى ؟

(جـ) هل إمكانية الشراء مستقلة عن نوع المشتري (ذكر أو أنثى) ؟

	لا يشترون	يشترون
إناث	100	80
ذكور	30	20

- ٤٦ - طائرة تطير يوميا بين مدينتي فإذا كان احتمال أن تقوم في ميعادها 0.8 احتمال أن يكون الجو جيد عندما تطير في موعدها هو 0.9 وعندما لا تطير في موعدها فإن احتمال أن يكون الجو رديئا هو 0.7 إذا ركب شخص الطائرة وكان الجو جيد ما هو احتمال أن الطيران يكون في ميعاده ؟

- ٤٧ - يفرض أن 1% من سكان مدينة ما يعانون من مرض ما . فإذا ظهر اختبار جديد للكشف عن المرض وأجرى على سكان المدينة . أعطى الاختبار نتيجة موجبة في 95% من الحالات التي عندها المرض . كما أعطى الاختبار نتيجة سالبة في 97% من الحالات التي ليس عندها المرض . اختر شخص بطريقة عشوائية وكانت نتيجة الكشف عنده موجبة ، ما احتمال أنه يعاني من المرض ؟

٤٨- مصنع ينتج ثلاثة أصناف من المصاييح بنسب 60% , 30% , 10% . التالف في الإنتاج هي 4% , 3% , 2% على التوالي اختبر إحدى أصناف الإنتاج واختبر منه مصباح أوجد:-

(أ) احتمال أن المصباح تالف .

(ب) إذا كان المصباح تالف أوجد احتمال أن يكون من إنتاج الصنف الأول .

٤٩ - تمثل الطالبات 30% من حجم الدارسين في كلية ما . يدرس 30% من الطلاب مادة الرياضيات وتدرس 20% من الطالبات مادة الرياضيات . إذا اختبر واحد من الكلية ووجد أنسه يدرس الرياضيات ما هو احتمال أن يكون المختار طالبة .

٥٠ - في دراسة ميدانية في إحدى الكليات وجد أن 7% من الذكور ، 2% من الإناث أطول من 1.7 مترا وأن 70% من الدارسين من الإناث . اختبر واحد عشوائيا ووجد أنه أطول من 1.7 فما احتمال أن تكون أنثى .

٥١ - إذا كان 20% من العاملين في شركة ما يحملون شهادات عليا وإذا كان 25% من الذين يحملون شهادات متوسطة يشغلون مناصب عليا . أيضا 75% من الذين يحملون شهادات عالية يشغلون مناصب عليا . فإذا اختبر أحد الأشخاص عشوائيا ما هو احتمال أن يحمل شهادة عليا إذا علم أنه يشغل منصب عالي .

٥٢ - لدينا ثلاث أوعية كما يلي :- الوعاء الأول به 4 كرات بيضاء ، والوعاء الثاني به ثلاث كرات حمراء و 4 بيضاء ، والوعاء الثالث به 2 حمراء و 3 بيضاء اختبر وعاء بطريقة عشوائية وسحبت منه كرة ووجد أن الكرة بيضاء فما احتمال أن تكون من الوعاء الأول ؟

٥٣ - يذهب رجل إلى عمله يوميا إما بسيارته أو بوسائل النقل العام . احتمال أن يركب سيارته هو 0.4 واحتمال أن يتأخر عن عمله إذا استخدم وسائل النقل العام هو 0.3 واحتمال أن يتأخر عن عمله إذا استخدم سيارته هو 0.1 فإذا ذهب إلى عمله متأخرا في يوما ما أوجد احتمال أن يكون قد أسفل سيارته .

الفصل الرابع

المتغيرات العشوائية

وتوزيعاتها الاحتمالية

Random Variables and their
Probability Distributions

Random Variable

(١-٤) المتغير العشوائي

تستخدم كلمة تجربة (كما ذكرنا سابقا) لأي إجراء نعلم مسبقا جميع النواتج الممكنة له وإن كنا لا نستطيع أن نتنبأ بأي من هذه النواتج سيتحقق فعلا. ربما لا يكون من الضروري دراسة فئة كل النواتج الممكنة (فراغ العينة) لتجربة إحصائية ولكن يكون اهتمامنا منصبا على قيم رقمية مرتبطة بهذه النواتج الممكنة. إن القيم الممكنة هذه هي ما نعر عنه بقيم المتغير العشوائي.

تعريف : الدالة المعرفة على فراغ العينة لتجربة ما والتي تخصص عددا حقيقيا لكل نقطة عينة تسمى المتغير العشوائي.

سوف نستخدم الرمز X ليمثل المتغير العشوائي، x لوحدة من قيمه.

مثال (١-٤) اختبرت بنوتان من نبات مزهر عشوائيا من كيس يحتوى على خمس بذور زهورها حمراء وثلاث بذور زهورها صفراء وذلك لاستخدامها في تجربة معينة. فراغ العينة يكون :

$$S = \{yy, ry, yr, rr\}$$

حيث r ترمز إلى البذرة التي زهورها حمراء، y ترمز إلى البذرة التي زهورها صفراء. بفرض أننا عرفنا الدالة X التي تمثل عدد البذور التي زهورها حمراء في العينة. هذه الدالة سوف تخصص عددا حقيقيا لكل نقطة عينة في فراغ العينة S المرافق لتجربتنا الإحصائية. في الجدول التالي نجد أن كل نقطة في فراغ العينة ارتبطت بعدد حقيقي واحد عن طريق الدالة X .

x	0	1	1	2
نقطة العينة	yy	ry	yr	rr

وعلى ذلك X متغير عشوائي يأخذ القيم 0, 1, 2

قد يحتوى فراغ العينة على عدد محدود من النقاط كما في المثال السابق، أو قد يكون فراغ العينة لإنهائي معدود **countable infinite sample space** وهو الفراغ الذي يحتوى على عدد لإنهائي من العناصر لكنه قابل للعد بمعنى أن هناك تقابل بين عناصره وفئة الأعداد الطبيعية، مثل عدد البكتيريا في لتر من الماء النقي أو عدد الفئران في فدان من القمح. يسمى فراغ العينة في هذه الحالة فراغ عينة منفصل (مقطع) **discrete sample space**. المتغير العشوائي المعرف على فراغ عينة منفصل يسمى متغير عشوائي منفصل (مقطع) **discrete random variable**. أيضا إذا كان فراغ العينة يحتوى على عدد لإنهائي من النقاط **infinite sample space** الغير معدودة مثل كل الأطوال الممكنة، الأوزان، درجات الحرارة،

العمر ١٠٠ الخ فإننا نقول أن فراغ العينة متصل (مستمر) continuous sample space
 المتغير العشوائي المعروف على فراغ عينة متصل يسمى المتغير العشوائي المتصل continuous
 random variable. في معظم التطبيقات المتغيرات العشوائية المنفصلة تمثل بيانات قابلة للعد ،
 مثل عدد الحوادث في السنة ، عدد الأخطاء في صفحة من قاموس ، عدد الفترات في فدان من
 القمح ١٠٠ الخ. أما المتغيرات العشوائية المتصلة فتتمثل ببيانات مقاسة .
 (٤-٧) التوزيعات الاحتمالية المنفصلة (المتقطعة)

Discrete Probability Distributions

كل قيمة من قيم المتغير العشوائي المنفصل يفرض لها احتمال ففي مثال (٤-١) تحسب
 الاحتمالات المختلفة لقيم المتغير العشوائي X الذي يمثل عدد البذور التي زهرها حراء في العينة
 (إذا كان الاختيار بدون إرجاع) كالآتي :-

$$P(X=0) = P(yy) = \left(\frac{3}{8}\right)\left(\frac{2}{7}\right) = \frac{6}{56} ,$$

$$P(X=1) = P(ry) + P(yr) \\ = \left(\frac{5}{8}\right)\left(\frac{3}{7}\right) + \left(\frac{3}{8}\right)\left(\frac{5}{7}\right) = \frac{30}{56} ,$$

$$P(X=2) = P(rr) = \left(\frac{5}{8}\right)\left(\frac{4}{7}\right) = \frac{20}{56} .$$

القيم المختلفة للمتغير العشوائي X مع احتمالاتها معطاة في الجدول التالي :-

x	0	1	2
P(X=x)	$\frac{6}{56}$	$\frac{30}{56}$	$\frac{20}{56}$

مجموع الاحتمالات في الجدول السابق تساوي الواحد الصحيح .

مثال (٤-٢) المطلوب الاحتمالات المختلفة لقيم المتغير العشوائي X في مثال (٤-١) إذا
 كان الاختيار بإرجاع :-

$$P(X=0) = P(yy) = \left(\frac{3}{8}\right)\left(\frac{3}{8}\right) = \frac{9}{64} , \quad \text{الحل .}$$

$$P(X=1) = P(ry) + P(yr) \\ = \left(\frac{5}{8}\right)\left(\frac{3}{8}\right) + \left(\frac{3}{8}\right)\left(\frac{5}{8}\right) = \frac{30}{64} ,$$

$$P(X=2) = P(rr) = \left(\frac{5}{8}\right)\left(\frac{5}{8}\right) = \frac{25}{64} .$$

القيم المختلفة للمتغير العشوائي X مع احتمالاتها معطاة في الجدول التالي :-

x	0	1	2
$P(X=x)$	$\frac{9}{64}$	$\frac{30}{64}$	$\frac{25}{64}$

عادة يفضل تحيل كل احتمالات المتغير العشوائي X بصيغة . هذه الصيغة من الضروري أن تكون دالة في القيم الرقمية x . سوف نرمز للدالة بوحدة من الصيغ $f(x), g(x), h(x)$. وعلى ذلك يمكن كتابة $f(x) = P(X=x)$ ، فعلى سبيل المثال $f(2) = P(X=2)$. الفئة من الأزواج المرتبة $(x, f(x))$ تسمى دالة الاحتمال **probability distribution** أو التوزيع الاحتمالي **probability distribution** للمتغير العشوائي X .

تعريف : كل جدول أو صيغة تعطي جميع القيم التي يأخذها متغير عشوائي منفصل، مع احتمال كل قيمة منها يسمى توزيع احتمالي منفصل .
مثال (٤-٣) إذا أُلقيت عملة مرة واحدة وكان X متغير عشوائي حيث $x=1$ إذا كانت النتيجة صورة و $x=0$ إذا كانت النتيجة كتابة . في هذه الحالة يمكن وضع صيغة للدالة $f(x)$ الخاصة بالمتغير العشوائي X على الشكل :-

$$f(x) = \frac{1}{2}, x = 0, 1.$$

مثال (٤-٤) أوجد صيغة التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل عدد الصور السقي تظهر عند إلقاء خمس عملات مرة واحدة ؟

الحل . عدد النقط في فراغ العينة سوف يكون 2^5 ، (الأحداث متساوية في إمكانية الحدوث) . المقام لجميع الاحتمالات سوف يكون 32 . لإيجاد عدد الطرق للحصول على x من الصور عند إلقاء 5 عملات مرة واحدة فإننا نحتاج لمعرفة العدد الكلي لنقاط العينة في التجربة والتي تعطي x صور و $n-x$ كتابة وهذا يساوي عدد الطرق لتبديل x من العناصر منها x من نوع (صورة) و $n-x$ من نوع آخر (كتابة) ، وهذا يحدث بطرق عددها $\binom{5}{x}$ حيث x تأخذ القيم 0,1,2,3,4,5 وعلى ذلك :-

$$f(x) = P(X=x) = \frac{\binom{5}{x}}{32}, x = 0,1,2,3,4,5.$$

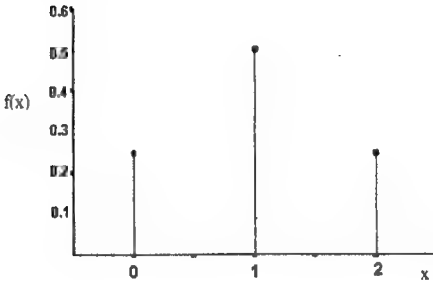
مثال (٤-٥) إذا كان X متغير عشوائي يمثل نواتج إلقاء زهرة نرد مرة واحدة فإن x تأخذ قيم من 1 إلى 6 . التوزيع الاحتمالي للمتغير العشوائي X هو :-

$$f(x) = \frac{1}{6}, \quad x = 1, 2, 3, 4, 5, 6.$$

مثال (٤-٦) إذا كان X متغيراً عشوائياً يمثل عدد الصورة التي تظهر عند إلقاء عملتين مرة واحدة فإن $x = 0, 1, 2$. التوزيع الاحتمالي للمتغير العشوائي X يمكن تمثيله بالجدول التالي :

x	0	1	2
$f(x)$	0.25	0.5	0.25

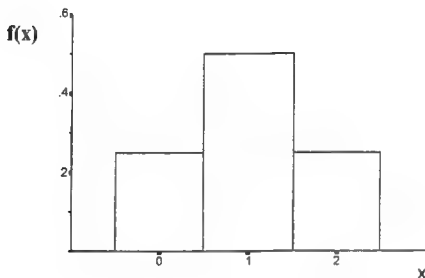
يمكن عرض هذا التوزيع بيانياً باستخدام طريقة الأعمدة bar chart كما في شكل (٤-١) .



شكل (٤-١)

حيث يمثل المحور الأفقي قيم x ويمثل المحور الرأسي قيم $f(x)$ فمثلاً عند $x = 0$ يقام عمود ارتفاعه يتناسب مع قيمة الدالة عند هذه النقطة وهو 0.25 وكذلك عند $x = 1$ يقام عمود ارتفاعه 0.5 وعند $x = 2$ يقام عمود ارتفاعه 0.25 وبخلاف هذه النقط فالدالة ليس لها وجود . كما يمكن تحويل شكل (٤-١) إلى ما يسمى بالدرج الاحتمالي probability histogram كما في شكل (٤-٢) وذلك بتحويل الأعمدة الموجودة إلى مستطيلات بحيث يكون ارتفاع كل مستطيل مساوياً لاحتمال قيمة x الواقعة في منتصف قاعدة المستطيل . وعلى

ذلك فإن $P(X = x)$ يساوى مساحة المستطيل الذي تقع x في منتصف قاعدته . هذا المفهوم لحساب الاحتمالات ضروري في التوزيع الاحتمالي المتصل .

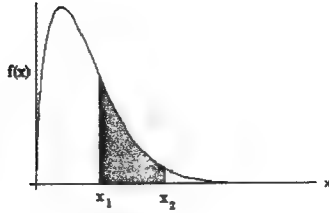


شكل (٤-٢)

(٤-٣) التوزيعات الاحتمالية المتصلة (المستمرة)

Continuous Probability Distributions

من صفات المتغير العشوائي المتصل أنه لا يمكن أن يكون هناك احتمال موجب مرافق لكل قيمة من قيم المتغير أى أن $P(X = x) = 0$ ولكن يكون هناك احتمال مرافق لكل فترة من فترات المتغير . ولهذا لا يمكن تمثيل التوزيع الاحتمالي للمتغير العشوائي المتصل بجدول ولكن نعبّر عنه بصيغة دالة $f(x)$ والتي تسمى دالة كثافة الاحتمال **probability density function** . التمثيل البياني للدالة $f(x)$ سوف يكون متصل ويأخذ أشكال كثيرة . واحد من هذه الأشكال موضح في شكل (٤-٣) ولابد أن تكون المساحة تحت منحنى الدالة $f(x)$ والمحددة بمحور x تساوى الواحد الصحيح . أيضا احتمال أن يأخذ المتغير العشوائي X قيمة بين $x = x_1$ و $x = x_2$ يساوى المساحة المظللة تحت دالة كثافة الاحتمال بين $x = x_1$ و $x = x_2$. المساحة بالضبط يمكن الحصول عليها باستخدام طرق التكامل . سوف يكون اهتمامنا فقط بدوال كثافة الاحتمال الشائعة الاستخدام في التجارب والتي تحسب المساحات تحت منحناها باستخدام الجداول الإحصائية . ولأن المساحات تمثل احتمالات والاحتمالات قيم موجبة، فإن دالة كثافة الاحتمال لابد أن تكون فوق منحنى x .

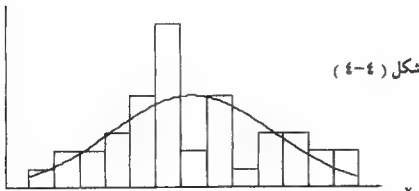


شكل (٣-٤)

تعريف : الدالة $f(x)$ تسمى دالة كثافة الاحتمال لمغير عشوائي متصل X إذا كانت المساحة الكلية تحت المنحنى والمحددة بمحور x تساوى الواحد الصحيح . أيضا المساحة تحت المنحنى بين أي قيمتين $x = x_1$ و $x = x_2$ تعطي احتمال أن المتغير العشوائي يقع بين $x = x_1$ و $x = x_2$

عادة في التجربة التي تحتوي على متغير متصل تكون $f(x)$ غير معلومة ويفرض معادلتها . الاختيار الصحيح للمعادلة يعتمد على توفر معلومات عن المتغير موضع الدراسة . المدرج يلقى الضوء على شكل $f(x)$. على سبيل المثال إذا كان الاهتمام بوزن مجموعة من الحيوانات فإننا نقوم بتوزيع المشاهدات (الأوزان x) إلى فئات ونقدر نسبة المشاهدات في كل فئة (فترة) والتي تسمى التكرارات النسبية **relative frequencies** للفئات المختلفة . هذه المعلومات يمكن تمثيلها بالمدرج ومنها نحصل على $f(x)$ التقريبية وذلك بتمهيد المنحنى كما في شكل (٤-٤) . شكل المنحنى يساعدنا في الاختيار المناسب للدالة $f(x)$.

التكرار النسبي



شكل (٤-٤)

Mathematical Expectation

(٤-٤) التوقع الرياضي

يمكن تسهيل لهم التوقع الرياضي بالمثال التالي : ليكن X متغير عشوائي يمثل عدد الصور التي تظهر عند إلقاء عمليتين مرة واحدة . وعلى ذلك X يأخذ القيم 0,1,2 باحتمالات 0.25, 0.5, 0.25 على التوالي . بفرض أن التجربة كررت بعدد كبير جدا من المرات، وليكن $n = 8000000$ ، نتوقع أن نلاحظ تقريبا 2 مليون للحادثة "عدم ظهور الصورة" و 4 مليون للحادثة "ظهور صورة و كتابة" و 2 مليون للحادثة "ظهور صورتين" . وعلى ذلك متوسط عدد الصور في الرمية الواحدة يساوى:

$$\begin{aligned} \frac{\text{Sum of observations}}{N} &= \frac{(0)(2000000) + (1)(4000000) + (2)(2000000)}{8000000} \\ &= \frac{(0)(2000000)}{8000000} + \frac{(1)(4000000)}{8000000} + \frac{(2)(2000000)}{8000000} \\ &= (0)\left(\frac{1}{4}\right) + (1)\left(\frac{1}{2}\right) + (2)\left(\frac{1}{4}\right) = 1. \end{aligned}$$

حيث عدد المشاهدات = sum of observations . يلاحظ أن الحد الأول يساوى $(0)f(0)$ والحد الثاني يساوى $(1)f(1)$ والحد الثالث يساوى $(2)f(2)$ وعلى ذلك يمكن تعريف القيمة المتوقعة للمتغير العشوائي X (متوسط التوزيع أو متوسط المجتمع) كالآتي :-

$$\mu = E(X) = (0)(.25) + (1)(0.5) + (2)(0.25) = 1.$$

تعريف : إذا كان X متغير عشوائي منفصل له التوزيع الاحتمالي التالي :-

x	x_1	x_2	...	x_n
$P(X=x)$	$f(x_1)$	$f(x_2)$...	$f(x_n)$

لأن التوقع الرياضي (القيمة المتوقعة أو متوسط المجتمع μ population mean) لمتغير عشوائي X هو :-

$$E(X) = \sum_{i=1}^n x_i f(x_i).$$

مثال (٧-٤) اختيرت عينة من 3 وحدات بطريقة عشوائية من صندوق به 12 وحدة بينها 3 معيبة أوجد القيمة المتوقعة لعدد الوحدات المعيبة .

الحل • يتكون فراغ العينة S من $\binom{12}{3} = 220$ عينة متساوية الاحتمال حجمها 3 • أيضا يوجد
 $\binom{9}{3}\binom{3}{0} = 84$ عينة ليست بها وحدات معيبة و $\binom{9}{2}\binom{3}{1} = 108$ عينة لها وحدة واحدة معيبة و
 $\binom{9}{1}\binom{3}{2} = 27$ عينة لها وحدتين معيبتين وعينة واحدة، $\binom{9}{0}\binom{3}{3} = 1$ ، لها ثلاث وحدات معيبة
 وبذلك يكون احتمال الحصول على 3,2,1,0 من الوحدات المعيبة على التوالي هو :-

$$\frac{1}{220}, \frac{27}{220}, \frac{108}{220}, \frac{84}{220},$$

أذن القيمة المتوقعة لعدد الوحدات المعيبة هي :-

$$E(X) = (0)f(0) + (1)f(1) + (2)f(2) + (3)f(3)$$

$$= (0)\left(\frac{84}{220}\right) + (1)\left(\frac{108}{220}\right) + (2)\left(\frac{27}{220}\right) + (3)\left(\frac{1}{220}\right) = \frac{165}{220}.$$

مثال (٤-٨) يحتوى صندوق على 4 كرات ، اثنين منهم مرقمين بالرقم 2 ، وكرة مرقمة بالرقم
 4 ، والأخيرة مرقمة بالرقم 8 سحب كرة من الصندوق أو جرد $E(X)$ ،
 الحل • قيم المتغير العشوائي سوف تكون $x = 2, 4, 8$ باحتمالات :-

$$f(2) = \frac{1}{2}, f(4) = \frac{1}{4}, f(8) = \frac{1}{4}.$$

وعلى ذلك القيمة المتوقعة للمتغير X أو متوسط المجتمع هو :-

$$\mu = E(X) = (2)\left(\frac{1}{2}\right) + (4)\left(\frac{1}{4}\right) + (8)\left(\frac{1}{4}\right) = 4.$$

في الجزء السابق تناولنا كيفية إيجاد القيمة المتوقعة لمتغير عشوائي • ليكن $h(X)$ دالة في
 متغير عشوائي (متغيراً عشوائياً جديداً يعتمد على X) • وعلى ذلك يمكن تقدير قيم الدالة $h(X)$
 بمعرفة قيم X • على سبيل المثال $h(X)$ قد تكون $(X - \mu)^2$ أو X^2 • وعلى ذلك إذا كان X
 يأخذ القيمة 4 فإن $h(X)$ تأخذ القيمة $h(4)$ وعلى ذلك يمكن كتابة :-
 $P[h(X) = h(4)] = P(X = 4) = f(4)$ • عموماً $P[h(X) = h(x)] = P(X = x) = f(x)$ •
 وعلى ذلك يمكن الحصول على القيمة المتوقعة للدالة $h(X)$ باستخدام الاحتمالات المعطاة للدالة
 $f(x)$ • على سبيل المثال إذا كان X متغيراً عشوائياً يمثل عدد الصور التي تظهر عند إلقاء عمليتين مرة
 واحدة وبفرض أننا نرغب في إيجاد القيمة المتوقعة للدالة $Y = X^2$ • التوزيع الاحتمالي للمتغير X
 وأيضاً للمتغير العشوائي Y في الجدول التالي :-

x	0	1	2
y	0	1	4
P(X=x)=P(Y=y)	0.25	0.5	0.25

$$E(Y) = (0)\left(\frac{1}{4}\right) + (1)\left(\frac{1}{2}\right) + (4)\left(\frac{1}{4}\right)$$

$$= \sum_{x=0}^2 x^2 f(x) = 1.5.$$

تعريف : إذا كان X متغير عشوائي منفصلا له التوزيع الاحتمالي التالي :-

x	x_1	x_2	...	x_n
P(X=x)	$f(x_1)$	$f(x_2)$...	$f(x_n)$

وإذا كان $h(X)$ دالة في X فإن $h(X)$ تمثل أيضا متغيرا عشوائيا والقيمة المتوقعة له هي :-

$$E[h(X)] = \sum_{i=1}^n h(x_i) f(x_i)$$

مثال (٩-٤) إذا كان X متغيرا عشوائيا بتوزيع احتمالي معطى في الجدول التالي أوجد القيمة المتوقعة للمتغير $-(X^2 - 1)$:-

الحل .

x	-1	0	1	2
P(X=x)	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$

$$E(X^2 - 1) = \sum_{x=-1}^2 (x^2 - 1) f(x)$$

$$= [(-1)^2 - 1]f(-1) + [(0)^2 - 1]f(0) + [(1)^2 - 1]f(1) + [(2)^2 - 1]f(2)$$

$$= (0)\left(\frac{1}{8}\right) + (-1)\left(\frac{1}{4}\right) + (0)\left(\frac{3}{8}\right) + (3)\left(\frac{1}{4}\right) = \frac{1}{2}.$$

مثال (١٠-٤) أوجد القيمة المتوقعة للدالة $(X - \mu)^2$ حيث أن X تمثل عدد الصور التي تظهر

عند إلقاء عملتين مرة واحدة .

الحل . قد أثبتنا من قبل أن $E(X) = 1$ ، وعلى ذلك فإن :-

$$\begin{aligned}
 E(X-\mu)^2 &= \sum_{x=0}^2 (x-\mu)^2 f(x) \\
 &= (0-1)^2 f(0) + (1-1)^2 f(1) + (2-1)^2 f(2) \\
 &= (1)(0.25) + (0)(0.5) + (1)(0.25) = 0.5.
 \end{aligned}$$

القيمة المتوقعة للالة $(X-\mu)^2$ تسمى التباين variance للمتغير العشوائي X ويرمز لها بالرمز σ^2 . يعرف الانحراف المعياري standard deviation للمتغير العشوائي X بأنه الجذر التربيعي لتباين X .

تعريف: التباين للمتغير العشوائي X هو :-

$$\sigma^2 = E(X-\mu)^2.$$

مثال (١١-٤) الجدول التالي يمثل التوزيع الاحتمالي للمتغير العشوائي X والذي يمثل عدد أجهزة الحاسب الآلي من 4 أجهزة والتي قد تعرض للتلف أثناء عملية الشحن إلى مركز أبحاث. أوجد التباين والانحراف المعياري للمتغير العشوائي X .

x	0	1	2	3	4
$P(X=x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

الحل. بما أن التباين للمتغير العشوائي X معرف بالصيغة $\sigma^2 = E(X-\mu)^2$. أولاً نحسب العدد المتوقع للأجهزة التالفة كالتالي :-

$$\begin{aligned}
 \mu = E(X) &= \sum_{x=0}^4 x f(x) \\
 &= 0 \cdot \left(\frac{1}{16}\right) + 1 \cdot \left(\frac{4}{16}\right) + 2 \cdot \left(\frac{6}{16}\right) + 3 \cdot \left(\frac{4}{16}\right) + 4 \cdot \left(\frac{1}{16}\right) \\
 &= \frac{4}{16} + \frac{12}{16} + \frac{12}{16} + \frac{4}{16} = \frac{32}{16} = 2.
 \end{aligned}$$

الآن :-

$$\begin{aligned}\sigma^2 &= E(X - \mu)^2 = \sum_{x=0}^4 (x - 2)^2 f(x) \\ &= (0 - 2)^2 \left(\frac{1}{16}\right) + (1 - 2)^2 \left(\frac{4}{16}\right) + (2 - 2)^2 \left(\frac{6}{16}\right) + (3 - 2)^2 \left(\frac{4}{16}\right) + (4 - 2)^2 \left(\frac{1}{16}\right) \\ &= (4) \left(\frac{1}{16}\right) + (1) \left(\frac{4}{16}\right) + (0) \left(\frac{6}{16}\right) + (1) \left(\frac{4}{16}\right) + (4) \left(\frac{1}{16}\right) = 1.\end{aligned}$$

• أيضا الانحراف المعياري للمتغير X هو $\sigma = \sqrt{\sigma^2} = 1$.

تعريف : العزم من الدرجة k (حيث k عدد صحيح موجب) حول نقطة الأصل للمتغير العشوائي X هو :-

$$\mu_k' = E(X^k).$$

والعزم من الدرجة k حول المتوسط هو :-

$$\mu_k = E(X - \mu)^k.$$

العزم من الدرجة الأولى حول الصفر يعطى متوسط المجتمع μ والعزم من الدرجة الثانية حول المتوسط يعطى التباين σ^2 • العزم بصفة عامة لهم استخدامات كثيرة في الإحصاء سوف نتناول بعضها في الفصل الخامس •

(٥-٤) بعض خواص القيم المتوقعة

Some Properties of Expected Values

في هذا البند سوف نقدم بعض النظريات والتي عن طريقها يمكن حساب توقعات بدلالة توقعات أخرى معروفة أو توقعات سهلة في الحساب • كل النتائج التالية صحيحة سواء للمتغيرات عشوائية منفصلة أو متصلة • البراهين التالية سوف تقتصر على المتغيرات العشوائية المنفصلة المحدودة •

نظرية (١-٤) بفرض أن X متغيراً عشوائياً و a, b ثابتين فإن :-

$$E(aX + b) = aE(X) + b.$$

البرهان :-

$$\begin{aligned}
E(aX + b) &= \sum_{i=1}^n (ax_i + b)f(x_i) \\
&= (ax_1 + b)f(x_1) + (ax_2 + b)f(x_2) + \dots + (ax_n + b)f(x_n) \\
&= a[x_1f(x_1) + x_2f(x_2) + \dots + x_nf(x_n)] \\
&\quad + b[f(x_1) + f(x_2) + \dots + f(x_n)] \\
&= a \sum_{i=1}^n x_i f(x_i) + b \sum_{i=1}^n f(x_i).
\end{aligned}$$

المجموع الأول من اليمين هو $E(X)$ والمجموع الثاني يساوي واحد صحيح ، وعلى ذلك فإن :-

$$E(aX + b) = aE(X) + b.$$

نتيجة (١) إذا كانت $a = 0$ ، فإن $E(b) = b$.

نتيجة (٢) إذا كانت $b = 0$ ، فإن $E(aX) = aE(X)$.

نظرية (٢-٤) التوقع الرياضي لمجموع دالتين (أو أكثر) في متغير عشوائي X يساوي مجموع القيم المتوقعة للدوال ، أي أن :-

$$E[g(X) + h(X)] = E[g(X)] + E[h(X)].$$

البرهان :-

$$\begin{aligned}
E[g(X) + h(X)] &= \sum_{i=1}^n [g(x_i) + h(x_i)]f(x_i) \\
&= [g(x_1) + h(x_1)]f(x_1) + [g(x_2) + h(x_2)]f(x_2) + \dots \\
&\quad + [g(x_n) + h(x_n)]f(x_n) \\
&= [g(x_1)f(x_1) + g(x_2)f(x_2) + \dots + g(x_n)f(x_n)] \\
&\quad + [h(x_1)f(x_1) + h(x_2)f(x_2) + \dots + h(x_n)f(x_n)] \\
&= \sum_{i=1}^n g(x_i)f(x_i) + \sum_{i=1}^n h(x_i)f(x_i) \\
&= E[g(X)] + E[h(X)].
\end{aligned}$$

نظرية (٣-٤) التباين للمتغير العشوائي X يعطى من الصيغة التالية :-

$$\sigma^2 = E(X)^2 - \mu^2.$$

البرهان :-

$$\begin{aligned}
 \sigma^2 &= E(X - \mu)^2 \\
 &= E(X^2 - 2\mu X + \mu^2) \\
 &= E(X^2) - 2\mu E(X) + E(\mu^2) \\
 &= E(X)^2 - \mu^2.
 \end{aligned}$$

- حيث أن $\mu = E(X)$ من التعريف و $E(\mu^2) = \mu^2$ من نظرية (١-٤) ونتيجة (١)
 مثال (١٢-٤) أوجد التباين للمتغير العشوائي X في مثال (١٠-٤)
 الحل . قد أثبتنا من قبل أن $E(X) = 1$ ، الآن :-

$$E(X^2) = \sum_{x=0}^2 x^2 f(x) = (0)(0.25) + (1)(0.5) + (2)^2(0.25) = 1.50.$$

وعلى ذلك :-

$$\begin{aligned}
 E(X - \mu)^2 &= E(X^2) - \mu^2 \\
 &= (1.5) - (1)^2 = 0.5.
 \end{aligned}$$

وهي نفس النتيجة التي حصلنا عليها من مثال (١٠-٤) .

- مثال (١٣-٤) أوجد التباين للمتغير العشوائي X في مثال (١١-٤) باستخدام الصيغة :-
 $E(X^2) - \mu^2.$

الحل . • $E(X) = 2$ تم الحصول عليه من مثال (١١-٤) الآن نوجد :-

$$\begin{aligned}
 E(X^2) &= \sum_{x=0}^4 x^2 f(x) \\
 &= (0)^2\left(\frac{1}{16}\right) + (1)^2\left(\frac{4}{16}\right) + (2)^2\left(\frac{6}{16}\right) + (3)^2\left(\frac{4}{16}\right) + (4)^2\left(\frac{1}{16}\right) \\
 &= \frac{4}{16} + \frac{24}{16} + \frac{36}{16} + \frac{16}{16} = \frac{80}{16} = 5.
 \end{aligned}$$

وعلى ذلك فإن التباين للمتغير العشوائي X هو :-

$$\begin{aligned}\sigma^2 &= E(X)^2 - \mu^2 \\ &= 5 - 2^2 = 1.\end{aligned}$$

وهي نفس النتيجة التي حصلنا عليها من مثال (١١-٤)

تعريف : ليكن X متغير عشوائي بـدالة كثافة احتمال $f(x)$ • البان لمطير عشوائي جديد $g(X)$ هو:-

$$\sigma_{g(X)}^2 = E[\{g(X) - \mu_{g(X)}\}^2].$$

نظرية (٤-٤) إذا كان X متغيرا عشوائيا و b ثابت ، فإن :-

$$\sigma_{X+b}^2 = \sigma_X^2 = \sigma^2.$$

البرهان :-

$$\sigma_{X+b}^2 = E[\{(X+b) - \mu_{X+b}\}^2].$$

الآن

$$\begin{aligned}\mu_{X+b} &= E(X+b) = E(X) + b \\ &= \mu + b,\end{aligned}$$

وذلك من نظرية (١-٤) وعلى ذلك :-

$$\begin{aligned}\sigma_{X+b}^2 &= E[(X+b - \mu - b)^2] \\ &= E[(X - \mu)^2] \\ &= \sigma^2.\end{aligned}$$

نظرية (٥-٤) إذا كان X متغير عشوائي و a ثابت، فإن :-

$$\sigma_{aX}^2 = a^2 \sigma_X^2 = a^2 \sigma^2.$$

البرهان :-

$$\sigma_{aX}^2 = E[\{aX - \mu_{aX}\}^2].$$

الآن :-

$$\mu_{aX} = E(aX) = aE(X) = a\mu,$$

وذلك من نظرية (١-٤) ، نتيجة (٢) وعلى ذلك :-

$$\begin{aligned}\sigma_{aX}^2 &= E[(aX - a\mu)^2] \\ &= a^2 E[(X - \mu)^2] \\ &= a^2 \sigma^2.\end{aligned}$$

(٤-٦) التوزيعات الاحتمالية الثنائية المنفصلة

Discrete Bivariate Distributions

بفرض أن لدينا متغيرين عشوائيين Y, X بتوزيع احتمالي $g(x), h(y)$ على التوالي . التوزيع الاحتمالي لوقوع Y, X في آن واحد عبارة عن صيغة دالة عادة يشار إليها بالرمز $f(x, y)$ وتسمى التوزيع الاحتمالي المشترك للمتغيرين Y, X . وعلى ذلك في حالة التوزيع المنفصل، فإن .
 $f(x, y) = P(X = x, Y = y)$ أي أن $f(x, y)$ تمثل احتمال وقوع Y, X في آن واحد على سبيل المثال إذا ألقينا زهرتي نرد مرة واحدة وإذا كانت X تمثل النقطة التي تظهر على السطح العلوي للزهرة الأولى و Y تمثل عدد النقاط التي تظهر على السطح العلوي للزهرة الثانية . فالتوزيع الاحتمالي المشترك للمتغيرين Y, X هو :-

$$P(X = x, Y = y) = f(x, y) = \frac{1}{36}, \quad x = 1, 2, 3, 4, 5, 6, \\ y = 1, 2, 3, 4, 5, 6.$$

تعريف : إذا كان Y, X متغيرين عشوائيين فما دالة التوزيع الاحتمالي المشترك $f(x, y)$ ، فإن هذه الدالة تحقق الشروط التالية :-

$$(١) \quad f(x, y) \geq 0 \quad \text{لجميع القيم } (x, y).$$

$$(ب) \quad \sum_y \sum_x f(x, y) = 1$$

من المثال السابق نجد أن $f(x, y) \geq 0$ لجميع القيم (x, y) حيث $f(x, y) = \frac{1}{36}$ لجميع القيم

$$\sum_y \sum_x f(x, y) = \sum_y \sum_x \frac{1}{36} = 1 \quad \text{و} \quad x = 1, 2, 3, 4, 5, 6 \quad ; \quad y = 1, 2, 3, 4, 5, 6, \quad (x, y)$$

إذا كان Y, X متغيرين عشوائيين منفصلين فما دالة التوزيع الاحتمالي المشترك $f(x, y)$ فإنه يمكن الحصول على التوزيع الاحتمالي للمتغير X عن حدة وتوزيع Y على حدة ويسمى التوزيع في هذه الحالة بالتوزيع الهامشي . وعلى ذلك تكون دالة التوزيع الاحتمالي الهامشي للمتغير X هي :

$$g(x) = \sum_y f(x, y).$$

وبالمثل دالة التوزيع الاحتمالي الهامشي للمتغير Y هي :-

$$h(y) = \sum_x f(x, y).$$

مثال (٤-١٤) اختبرت عينة من شخصين لإجراء اختبار معين عليهم من بين مجموعة مكونة من أربعة غير مدخنين وأثنين مدخنين. إذا كان X و Y معرفتان كالتالي $x=0$ إذا كان الشخص الأول غير مدخن و $x=1$ إذا كان الشخص الأول مدخناً. أيضاً $y=0$ إذا كان الشخص الثاني غير مدخن و $y=1$ إذا كان الشخص الثاني مدخناً. فإن التوزيع الاحتمالي المشترك للمتغيرين X و Y في الجدول التالي. المطلوب إيجاد التوزيع الهامشي لكل من X و Y .

x	0	1	$h(y)$
y			
0	$\frac{6}{15}$	$\frac{4}{15}$	$\frac{10}{15}$
1	$\frac{4}{15}$	$\frac{1}{15}$	$\frac{5}{15}$
$g(x)$	$\frac{10}{15}$	$\frac{5}{15}$	

الحل: أولاً، لحساب دالة التوزيع الهامشي للمتغير X نجمع الأعمدة كما يلي :-

$$g(0) = \sum_{y=0}^1 f(0, y) = f(0,0) + f(0,1) = \frac{10}{15},$$

$$g(1) = \sum_{y=0}^1 f(1, y) = f(1,0) + f(1,1) = \frac{5}{15}.$$

ثانياً، لحساب دالة التوزيع الهامشي للمتغير Y نجمع الصفوف كما يلي :-

$$h(0) = \sum_{x=0}^1 f(x, 0) = f(0,0) + f(1,0) = \frac{10}{15},$$

$$h(1) = \sum_{x=0}^1 f(x, 1) = f(0,1) + f(1,1) = \frac{5}{15}.$$

تعريف : إذا كان X, Y متغيرين عشوائيين بدالة كثافة احتمالية مشتركة، فإن الدالة

الاحتمالية المشروطة للمتغير Y بشرط أن $X = x$ تعرف بالصيغة :-

$$f_{Y|X}(y|x) = \frac{f(x,y)}{g(x)}.$$

• لقيم x بحيث أن $g(x) > 0$

وبنفس الشكل فإن الدالة الاحتمالية المشروطة للمتغير X بشرط أن $Y = y$ تعرف بالصيغة :-

$$f_{X|Y}(x|y) = \frac{f(x,y)}{h(y)}.$$

• لقيم y بحيث أن $h(y) > 0$

مثال (١٥-٤) أوجد $f_{Y|X}(y|0)$, $f_{X|Y}(x|1)$ للبيانات في مثال (١٤-٤) :-

الحل • أولاً، الدالة $f_{X|Y}(x|1)$ يمكن إيجادها كالآتي :-

$$f_{X|Y}(0|1) = \frac{f(0,1)}{h(1)} = \frac{4}{5},$$

$$f_{X|Y}(1|1) = \frac{f(1,1)}{h(1)} = \frac{1}{5}.$$

ثانياً، الدالة $f_{Y|X}(y|0)$ يمكن إيجادها كالآتي :-

$$f_{Y|X}(0|0) = \frac{f(0,0)}{g(0)} = \frac{6}{10},$$

$$f_{Y|X}(1|0) = \frac{f(0,1)}{g(0)} = \frac{4}{10}.$$

تعريف : يكون المتغيرين العشوائيين X و Y مستقلين إذا كان :-

$$f(x,y) = g(x)h(y)$$

لكل قيم (x,y) •

مثال (١٦-٤) في المثال (١٤-٤) هل Y, X مستقلين ؟

الحل • Y, X غير مستقلين لأنه بالنظر إلى الجدول المعروض في مثال (١٤-٤) نجد :-

$$f(x,y) \neq g(x)h(y), x=0,1, y=0,1.$$

$$\text{فعلى سبيل المثال } f(0,1) \neq g(0)h(1) \text{ حيث } f(0,1) = \frac{4}{15}, g(0) = \frac{10}{15}, h(1) = \frac{5}{15}$$

نظرية (١٧-٤) إذا كان Y, X متغيرين عشوائيين فإن :-

$$E(X \pm Y) = E(X) \pm E(Y).$$

فعلى سبيل المثال إذا أقيمت زهرة نرد مرتين وكانت X تمثل عدد النقاط الذي تظهر على السطح

العلوي في المرة الأولى و Y عدد النقاط التي تظهر على السطح العلوي في المرة الثانية فإن

$Y + X$ يمثل مجموع العددين اللذان يظهران على السطح العلوي للنرد عند إلقائها مرتين •

نظرية (١٧-٤) إذا كان Y, X متغيرين عشوائيين مستقلين فإن :-

$$E(XY) = E(X)E(Y).$$

على سبيل المثال عند إلقاء نردتين مرة واحدة فإن XY تمثل حاصل الضرب للعديدين الظاهريين على النردتين .

نظرية (٨-٤) يفرض أن المتغيرين العشوائيين Y, X مستقلين , فإن :-

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2.$$

البرهان :-

$$\sigma_{X+Y}^2 = E[\{(X+Y) - \mu_{X+Y}\}^2].$$

الآن :-

$$\mu_{X+Y} = E(X+Y) = E(X) + E(Y) = \mu_X + \mu_Y,$$

وذلك من نظرية (٦-٤) وعلى ذلك :-

$$\begin{aligned}\sigma_{X+Y}^2 &= E[\{(X+Y) - (\mu_X + \mu_Y)\}^2] \\ &= E[\{(X - \mu_X) + (Y - \mu_Y)\}^2] \\ &= E[(X - \mu_X)^2] + E[(Y - \mu_Y)^2] \\ &\quad + 2E[(X - \mu_X)(Y - \mu_Y)].\end{aligned}$$

الحدين الأولين يمثلان على التوالي σ_X^2, σ_Y^2 . المطلوب إثبات أن الحد الأخير يساوى صفر .

وعلى ذلك :-

$$\begin{aligned}E[(X - \mu_X)(Y - \mu_Y)] &= E(XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y) \\ &= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X \mu_Y \\ &= E(XY) - \mu_X \mu_Y = 0,\end{aligned}$$

وذلك لان $E(XY) = E(X)E(Y)$ للمتغيرات المستقلة . وعلى ذلك :-

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2.$$

نتيجة (٣) يفرض أن Y, X متغيرين عشوائيين مستقلين، فإن :-

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2.$$

النتيجة نحصل عليها بوضع $X - Y$ على الشكل $X + (-Y)$ وعلى ذلك :-

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_{(-Y)}^2.$$

من نظرية (٥-٤) نعلم أن $\sigma_{(-Y)}^2 = (-1)^2 \sigma_Y^2 = \sigma_Y^2$ وعلى ذلك :-

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2.$$

تعريف : القيمة المتوقعة للدالة $(X - \mu_X)(Y - \mu_Y)$ تعرف بالتغاير بين المتغيرين Y, X ويرمز لها بالرمز $Cov(X, Y)$ أي أن :-

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

وهو يقيس درجة الترابط بين المتغيرين . بعض خواص التغاير معطاة في النظريات التالية .

نظرية (٩-٤) إذا كانت Y, X متغيرين عشوائيين و a, b ثابتين فإن :-

$$\begin{aligned} Cov(aX, bY) &= abCov(X, Y), \\ Cov(X+a, Y+b) &= Cov(X, Y), \\ Cov(X, aX+b) &= a\sigma_X^2. \end{aligned}$$

مثال (١٧-٤) إذا كان Y, X متغيرين عشوائيين مستقلين حيث :-

$$\sigma_Y^2 = 16, \sigma_X^2 = 4, E(Y) = 3, E(X) = 2$$

أوجد : (١) $E(5X - Y)$ ، (ب) σ_{X-Y}^2 ، (جـ) $Cov(3X+2, Y)$

$$(د) Cov(X, 5X-2)$$

$$\text{الحل. (١) } E(5X - Y) = 5E(X) - E(Y) = (5)(2) - 3 = 7$$

(ب) وحيث Y, X مستقلين فإن :-

$$\begin{aligned} \sigma_{X-Y}^2 &= \sigma_X^2 + \sigma_Y^2 \\ &= 4 + 16 = 20. \end{aligned}$$

$$(جـ) Cov(3X+2, Y) = 3Cov(X, Y) = (3)(0) = 0$$

$$(د) Cov(X, 5X-2) = 5Cov(X, X) = 5\sigma_X^2 = (5)(4) = 20$$

نظرية (١٠-٤) إذا كان Y, X متغيرين عشوائيين، فإن :-

$$Cov(X, Y) = E(XY) - E(X)E(Y).$$

كما أن $Cov(X, Y) = 0$ إذا كان Y, X مستقلين بينما العكس ، عموماً ، غير صحيح بمعنى أنه بالإمكان أن يكون $Cov(X, Y) = 0$ ولكن Y, X غير مستقلين .

نظرية (١١-٤) إذا كان Y, X متغيرين عشوائيين بدالة كثافة احتمال مشتركة $f(x, y)$ ، فإن :

$$\sigma_{X \pm Y}^2 = \sigma_X^2 + \sigma_Y^2 \pm 2Cov(X, Y)$$

و $\sigma_{X \pm Y}^2 = \sigma_X^2 + \sigma_Y^2$ إذا كان Y, X مستقلين .

مثال (١٨-٤) إذا كان Y, X متغيرين عشوائيين بدالة كثافة احتمال :

$$f(x,y) = \frac{1}{4}$$

حيث $(x,y) = (0,1), (1,0), (0,-1), (-1,0)$ و $g(\pm 1) = \frac{1}{4}, g(0) = \frac{1}{2}$ ودالة كثافة الاحتمال للمتغير Y نفس دالة كثافة الاحتمال للمتغير X وبما أن $E(XY) = 0$ فإن :

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0.$$

ولكن $f(1,0) \neq g(1)h(0)$ ، وعلى ذلك Y, X متغيرين غير مستقلين. عموماً يمكن القول أن Y, X غير مستقلين إذا كان $\text{Cov}(X, Y) \neq 0$ ، ولكن $\text{Cov}(X, Y) = 0$ لا يعنى أن المتغيرين Y, X مستقلين.

تعريف : إذا كان Y, X متغيرين عشوائيين يتباين σ_X^2, σ_Y^2 وتباين $\text{Cov}(X, Y)$ ، فإن معامل الارتباط بين Y, X هو :-

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

يقال للمتغيرين العشوائيين Y, X أنهم غير مرتبطين إذا كان $\rho = 0$ ، غير ذلك يقال أنهما مرتبطين. نظرية (٤-١٢) إذا كان ρ معامل الارتباط بين المتغيرين Y, X فإن :-

$$-1 \leq \rho \leq 1.$$

مثال (٤-١٩) إذا كان Y, X متغيرين عشوائيين بدالة كثافة احتمال مشتركة :-

$$f(x,y) = \frac{4}{5xy}, x=1,2 \text{ and } y=2,3.$$

أوجد : (١) معامل الارتباط بين Y, X (ب) هل Y, X مستقلين أم لا ؟ .
الحل . (١) التوزيع الاحتمالي المشترك للمتغيرين Y, X في الجدول التالي :

$x \backslash y$	1	2
2	$\frac{12}{30}$	$\frac{6}{30}$
3	$\frac{8}{30}$	$\frac{4}{30}$

وعلى ذلك :-

$$E(X) = (1)\left(\frac{20}{30}\right) + (2)\left(\frac{10}{30}\right) = \frac{4}{3},$$

$$E(Y) = (2)\left(\frac{18}{30}\right) + (3)\left(\frac{12}{30}\right) = \frac{12}{5},$$

$$E(X^2) = (1)^2\left(\frac{20}{30}\right) + (2)^2\left(\frac{10}{30}\right) = 2,$$

$$E(Y^2) = (2)^2\left(\frac{18}{30}\right) + (3)^2\left(\frac{12}{30}\right) = 6,$$

$$\sigma_X^2 = E(X^2) - \mu_X^2 = 2 - \left(\frac{4}{3}\right)^2 = \frac{2}{9},$$

$$\sigma_Y^2 = E(Y^2) - \mu_Y^2 = 6 - \left(\frac{12}{5}\right)^2 = \frac{6}{25},$$

$$E(XY) = (1)(2)\left(\frac{12}{30}\right) + (2)(2)\left(\frac{6}{30}\right) + (1)(3)\left(\frac{8}{30}\right) + (2)(3)\left(\frac{4}{30}\right) = \frac{48}{15},$$

وعلى ذلك :-

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{48}{15} - \left(\frac{4}{3}\right)\left(\frac{12}{5}\right) = 0.$$

الارتباط بين Y, X هو :-

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{0}{\sqrt{\left(\frac{2}{9}\right)\left(\frac{6}{25}\right)}} = 0.$$

(ب) Y, X غير مستقلين لان $f(x, y) \neq g(x)h(y)$ لجميع قيم (x, y) .

تمارين

١ - صف المتغيرات العشوائية التالية إلى منفصلة ومتصلة :-

- (أ) الزمن اللازم لوصول طائرة (ب) الزمن اللازم لإنهاء امتحان (ج) عدد المصايح الثالثة في صندوق يحتوي على 5 مصايح (د) عدد الأخطاء التي يعرض لها شخص ما عند كتابة خطاب على الآلة الكاتبة (ز) كمية اللبن الحليب التي تدرها بقرة في العام (ر) عدد البيض الذي تضعه دجاجة في الشهر .

٢ - ألقيت عملة متحيزة ثلاث مرات بحيث أن فرصة ظهور الصورة ضعف فرصة ظهور الكتابة .
أوجد التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل عدد الصور التي تظهر على السطح العلوي للعملة .

٣ - الجدول التالي يمثل التوزيع الاحتمالي للمتغير عشوائي X :-

x	0	1	2
$P(X=x)$	0.4	0.2	

(١) ما هي قيمة $P(X=0)$ ؟ (ب) أوجد القيمة المتوقعة للمتغير X (ج) $P(X>1)$.

٤ - إذا ألقيت زهرتي نرد مرة واحدة أوجد :-

(١) التوزيع الاحتمالي للمتغير العشوائي Y الذي يمثل مجموع النقط التي تظهر على السطح العلوي للتردين ومثله بيانياً .

(ب) التوقع والتباين للمتغير Y .

(ج) التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل القيمة المطلقة للفرق بين مجموع النقط التي تظهر على السطح العلوي للتردين .

(د) التوقع والتباين للمتغير X .

٥ - إذا كان التوزيع الاحتمالي للزيادة في سعر سلعة ما في خلال سنة قادمة محددة كما في الجدول

التالي حيث $X=0$ تعني عدم وجود زيادة و $X=1$ زيادة أقل من 3% و $X=2$ زيادة من 3% إلى

6% و $X=3$ زيادة أكثر من 6% .

x	0	1	2	3
$P(X=x)$	0.1	0.1	0.5	0.3

أوجد : التوقع والتباين للمتغير X .

٦ - أوجد الصيغة الاحتمالية للمتغير X الذي يمثل عدد الصور التي تظهر على السطح العلوي عند

إلقاء عملة متزنة سبع مرات وأيضاً التوقع والتباين للمتغير X .

٧ - بفرض أن شركة شحن اشترت سيارة كبيرة بمبلغ 15000 دولار . إذا فقدت السيارة سواء

بالسرقة أو بمحادنة فإن ذلك يمثل فقد كلي . الفرصة في الفقد 0.002 ، أوجد القيمة المتوقعة للفقد

(المتغير العشوائي هنا يأخذ القيمة 0 لعدم الفقد والقيمة 15000 للفقد) .

٨ - أوجد القيمة المتوقعة لعدد الرجال الذين يتم اختيارهم لمهمة علمية من 3 أشخاص مسن بين 5

رجال وسيدتين .

٩ - إذا كان X متغيراً عشوائياً يمثل عدد أجهزة الحاسب الآلي التي تتعرض التلف من بين خمسة أجهزة وذلك أثناء توصيلها إلى مركز أبحاث . بفرض أن احتمال التلف 0.25 . أيضاً بفرض أن كسل جهاز مستقل عن الآخر في التلف أو عدم التلف أوجد القيمة المتوقعة للأجهزة التالفة .

١٠ - إذا كانت المبيعات من سلعة ما في الساعة هي $20, 21, 22$ عبوة باحتمال $0.2, 0.5, 0.3$ على التوالي . أوجد القيمة المتوقعة والتباين لعدد العبوات المباعة في الساعة .

١١ - احتمال أن يحصل لاعب كرة التنس على هدف في أى مباراة يلعبها هو 0.3 . أوجد القيمة المتوقعة لعدد الأهداف التي يكسبها في خمس مباريات قادمة .

١٢ - بفرض أن إيرادات متجر في اليوم تمثل بالمتغير العشوائي X ، والذي له التوزيع الاحتمالي

التالي :-

x	0	10	12	16	18
$P(X=x)$	0.1	0.2	0.2	0.3	0.2

أوجد التوقع والتباين .

١٣ - إذا كانت دالة كثافة الاحتمال لمغير عشوائي X هي :-

$$f(x) = \frac{1}{5}, x = 1, 2, 3, 4, 5.$$

أوجد احتمال أن X : (١) عدد زوجي - عدد فردي (ب) التوقع والتباين للمتغير X .

١٤ - يقوم بائع بتوصيل نوعين من المنظفات (B, A) إلى المنزل . المكسب من المنظف A و B

على التوالي هو $5, 10$ جنيهات للعبوة . الفرصة لبيع المنظف A هي 2 من 10 جولات والفرصة لبيع المنظف B هي 3 من 10 جولات والفرصة لعدم البيع هي 5 من 10 جولات . أوجد القيمة المتوقعة للمكسب في الجولة الواحدة .

١٥ - لدى محل للرياضة 80 عبوة تحتوي كل عبوة على كرات تنس ذات لون واحد ، إما صفراء أو خضراء . إذا كان عدد العلب التي تحتوي على كرات صفراء 30 . سمحت عبوة من 10 علب . أوجد : (١) التوزيع الاحتمالي لعدد العلب التي تحتوي على كرات صفراء (ب) القيمة المتوقعة لعدد العلب التي تحتوي على كرات صفراء .

١٦ - أى من الدوال التالية تمثل توزيع احتمالي :-

$$f(x) = \frac{x+2}{15}, x = -2, -1, 0, 1, 2 \quad (ب) \quad f(x) = \frac{2x}{5}, x = 0, 1, 2 \quad (١)$$

$$f(x) = \frac{x}{3}, x = -1, 0, 1, 2 \quad (\text{د}) \quad f(x) = \frac{x^2 - 1}{49}, x = 0, 1, 2, 3, 4, 5 \quad (\text{جـ})$$

$$f(x) = \frac{x^2 - x}{5}, x = 0, 1, 2, 3 \quad (\text{ز})$$

١٧ - في كل من الدوال التالية ، عين الثابت k بحيث تكون $f(x)$ دالة كثافة احتمالية :-

$$f(x) = \frac{k}{x}, x = 1, 2, 3 \quad (\text{ب}) \quad f(x) = \frac{k}{x^2}, x = 1, 2 \quad (\text{أ})$$

$$f(x) = k\left(\frac{1}{2}\right)^x, x = 1, 2, 3 \quad (\text{د}) \quad f(x) = kx, x = 0, 1, 2 \quad (\text{جـ})$$

$$f(x) = k\left[\left(\frac{1}{2}\right)^x - \frac{1}{2}\right], x = 0, 1, 2 \quad (\text{ز})$$

$$f(x) = k(8 - x), x = 0, 1, 2, 3, 4, 5 \quad (\text{ر})$$

١٨ - إذا كان X متغيراً عشوائياً يمثل الزمن بالتواني الذي يستغرقه حاسب في تنفيذ برنامج محاسب .
إذا كان التوزيع الاحتمالي للمتغير X هو :-

$$f(x) = \frac{x}{21}, x = 1, 2, 3, 4, 5, 6.$$

(أ) أثبت أن $f(x)$ دالة كثافة احتمالية (ب) ما هو احتمال أن الزمن الذي يستغرقه الحاسب :

بالضبط 4 ثواني في التنفيذ - على الأقل 3 ثواني وليس أكثر من 5 ثواني - أكثر من 5 ثواني .

١٩ - إذا كان التوزيع الاحتمالي لعدد العملاء الذين يترددون على شركة ما في اليوم هو :-

x	$P(X=x)$	x	$P(X=x)$
5	0.02	30	0.10
10	0.05	35	0.10
15	0.15	40	0.09
20	0.20	45	0.04
25	0.25		

أوجد القيمة المتوقعة لعدد العملاء في يوم محدد .

٢٠ - أى من الدوال التالية تمثل توزيع احتمالي :-

$$f(x) = \frac{2}{x}, x = 3, 4, 5 \quad (\text{ب}) \quad f(x) = x, x = \frac{1}{6}, \frac{1}{3}, \frac{1}{2} \quad (\text{أ})$$

$$f(x) = \frac{x-3}{9}, x = 3, 4, 5, 6, 7 \quad (\text{د}) \quad f(x) = \frac{x^2}{4}, x = 1, 2, 3 \quad (\text{جـ})$$

$$f(x) = \frac{x^2 - 1}{50}, x = 2, 3, 4, 5 \quad (د) \quad f(x) = \frac{x}{3}, x = \frac{1}{2}, 1, \frac{3}{2} \quad (ز)$$

$$f(x) = \frac{1}{x}, x = 2, 3, 4, 5 \quad (هـ)$$

- ٢١ - إذا كان التوزيع الاحتمالي لمغير X هو :-

$$f(x) = \binom{3}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, x = 0, 1, 2, 3.$$

أوجد : (أ) توقع وتباين المغير العشوائي X (ب) $E[(x - E(x))^2]$

$$E[(2x+1)^2] \quad (جـ) \quad E(2x^2 + 6) \quad (د)$$

- ٢٢ - الجدول الآتي يعطى الدالة الاحتمالية للمغيرين Y, X احسب معامل الارتباط وأوجد

$$\text{Cov}(X, 3X - 7), \text{Cov}(2X, Y), \sigma_{3X-2Y}^2, E(7X - 2Y)$$

y	x	0	1	$h(y)$
0		$\frac{6}{15}$	$\frac{4}{15}$	$\frac{10}{15}$
1		$\frac{4}{15}$	$\frac{1}{15}$	$\frac{5}{15}$
$g(x)$		$\frac{10}{15}$	$\frac{5}{15}$	

- ٢٣ - الجدول التالي يمثل التوزيع الاحتمالي لعدد الوحدات المعيبة في الإنتاج اليومي لأحد المصانع .

x	0	1	2	3	4	5
$P(X=x)$	0.884	0.1	0.01	0.003	0.002	0.001

المطلوب : (أ) تمثيل التوزيع بيانياً (ب) التباين والانحراف المعياري لعدد الوحدات المعيبة .

- ٢٤ - وعاء يحتوي على 40 كرة مرقمة من 1 إلى 40 فإذا تقرر اختيار كرة من الوعاء أذكر المتغير

العشوائي X الذي يمثل رقم الكرة المختارة من الصندوق وأوجد توقعة وتباينه .

الفصل الخامس

عرض ووصف البيانات

Presentation and Description of
Data

Populations and Samples

(١-٥) المجتمعات والعينات

تسجل نتيجة كل تجربة إحصائية ، كما ذكرنا في الفصل الثالث ، إما بقيمة رقمية أو تمثيل وصفي . فعلى سبيل المثال عند إلقاء زهرة نرد مرة واحدة وإذا كان الإهتمام بعدد النقاط التي تظهر على السطح العلوي للنرد فإننا نسجل قيمة رقمية . بينما عند سؤال مجموعة من العاملين في هيئة ما عن الحالة الاجتماعية لكل منهم ، فإن التمثيل الوصفي يكون أكثر فائدة . فالحالة الاجتماعية لأي شخص إما أعزب أو متزوج أو مطلق أو أرمل . عادة يهتم الإحصائي بالقيم الرقمية لذلك فإن التمثيل الوصفي يمكن تحويله إلى قيم عددية . فعلى سبيل المثال عند تصنيف مجموعة من الأشخاص حسب الحالة الاجتماعية فإنه يمكن تخصيص الرقم 1 للأعزب والرقم 2 للمتزوج والرقم 3 للمطلق والرقم 4 للأرمل . القيمة التي تسجل من نتيجة تجربة إحصائية تسمى بيان أو مشاهدة (مقياس) كما ذكرنا في الفصل الثالث . عندما يقوم باحث بتصنيف العاملين في شركة ما حسب الحالة الاجتماعية ، في هذه الحالة يكون لديه عدد محدود من المشاهدات . بينما عند إلقاء زهرة نرد عدد لا نهائي من المرات وتبجيل عدد النقاط التي تظهر في كل مرة فإننا نحصل على فئة لا نهائية من القيم . كل المشاهدات تحت الدراسة ، سواء كانت محدودة أو غير محدودة ، تسمى مجتمع **population** . في السنوات الماضية كانت كلمة مجتمع تشير إلى مشاهدات من دراسات إحصائية تشمل أشخاص . أما الآن فإن الإحصائي يستعمل هذه الكلمة لتشير إلى مشاهدات عن أي شيء موضع إهتمامه سواء مجموعة من الأشخاص ، حيوانات ، نباتات ... الخ .

تعريف : يتكون المجتمع من كل الأشياء التي نهتم بها .

عدد المشاهدات في المجتمع تسمى حجم المجتمع وعادة يرمز لحجم المجتمع بالرمز N ، وفي هذه الحالة نقول أن المجتمع محدود . فعلى سبيل المثال عند تصنيف 500 شخصاً في شركة ما حسب الحالة الاجتماعية ، فإننا نقول أن المجتمع محدود وحجمه $N=500$. الأسئلة والأوزان والدخل السنوي لمجموعة من الأشخاص أمثلة لمجتمعات محدودة . في كل حالة العدد الكلي للمشاهدات رقم محدود . في بعض الأحيان يكون حجم المجتمع غير محدود ، مثل مجتمع كرات الدم البيضاء التي تسري في دم إنسان . أيضاً المشاهدات التي نحصل عليها من قياس الضغط الجوي كل يوم من الماضي إلى المستقبل تمثل مجتمع غير محدود .

كل مشاهدة في المجتمع تمثل قيمة من قيم المتغير العشوائي X . على سبيل المثال عند إلقاء زهرة نرد عدد لا نهائي من المرات وإذا كان X يمثل عدد النقاط التي تظهر على النرد كل مرة ،

أى أن $x=1,2,3,4,5,6$ ، فإن كل مشاهدة في المجتمع تمثل قيمة من قسم المتغير العشوائي X .

لمزيد من التوضيح بفرض أن المجتمع التالي يمثل عدد مرات الغياب في السنة المسجلة لكل من عشرين طالبا في كلية ما: $0,0,7,8,9,9,9,8,6,5,4,3,2,1,1,6,6,7,8$. وعلى ذلك X متغير عشوائي يأخذ القيم $0,1,2,3,4,5,6,7,8,9$. قيم X مع الإحتمالات المقابلة لها تعرف التوزيع الإحتمالي للمتغير العشوائي X . وبما أن الإحتمالات تمثل تكرارات نسبية، فإن المجتمع أصبح معروف تماما من توزيعه الاحتمالي . إن دراسة خواص المتغير العشوائي تكافئ دراسة خواص المجتمع. على سبيل المثال القيمة المتوقعة للمتغير X هي :-

$$E(X) = (0)\left(\frac{2}{20}\right) + (1)\left(\frac{2}{20}\right) + (2)\left(\frac{1}{20}\right) + (3)\left(\frac{1}{20}\right) + (4)\left(\frac{1}{20}\right) + (5)\left(\frac{1}{20}\right) +$$

$$(6)\left(\frac{3}{20}\right) + (7)\left(\frac{2}{20}\right) + (8)\left(\frac{4}{20}\right) + (9)\left(\frac{3}{20}\right) = \frac{107}{20} = 5.35.$$

والتي تساوى الوسط الحسابي للمجتمع μ وذلك من المعادلة التالية :-

$$\mu = \frac{\sum_{i=1}^N x_i}{N} = \frac{107}{20} = 5.35.$$

تعريف : القيم العددية التي توصف المجتمع تسمى معالم.

يهتم الباحث بالوصول إلى إستنتاجات تخص معالم المجتمع، ولكن عادة يكون من المستحيل أو غير عملي ملاحظة كل قيم الفئة المثلة للمجتمع. وعلى ذلك لابد من الإعتماد على فئة جزئية من قيم المجتمع لمساعدنا في الوصول إلى إستدلالات عن المعالم، وهذا يأخذنا إلى نظرية المعاينة *theory of sampling* .

تعريف : العينة *sample* هي فئة جزئية من المجتمع.

حتى يكون الإستدلال صحيح لابد من فهم العلاقة بين المجتمع والعينة. من المؤكد أن العينة سوف تمثل المجتمع لذلك لابد أن تكون غير متحيزة *unbiased* أى عينة عشوائية *random sample* .

تعريف : العينة العشوائية من الحجم n هي عينة تختار بحيث أن كل فئة جزئية حجمها n من مشاهدات المجتمع لها نفس الإحتمال في الاختيار.

قد نرغب في الوصول إلى إستنتاجات تخص نسبة الأشخاص المدخنين في بلد ما . في بعض الأحيان يكون من الصعوبة سؤال كل شخص في هذا البلد وحساب المعلمة التي تمثل نسبة

المُدخَين الحقيقية. بدلا من ذلك نختار عينة عشوائية كبيرة ونحسب النسبة من العينة. هذه القيمة تستخدم في عمل بعض الاستدلال الذي يخص النسبة الحقيقية. القيمة المحسوبة من العينة تسمى الإحصاء statistic . وبما أن عينات عشوائية كثيرة يمكن إختيارها من نفس المجتمع فإننا نتوقع أن يختلف الإحصاء من عينة إلى أخرى، وعلى ذلك يعتبر الإحصاء متغير عشوائي .

تعريف : الإحصاء متغير عشوائي يعتمد فقط على قيم العينة المختارة .

عادة، يمثل قيمة أى إحصاء بحرف من الحروف اللاتينية الصغيرة . على سبيل المثال إذا أخذنا عينة عشوائية من الحجم $n=3$ من المجتمع الذي يمثل عدد مرات الغياب لـ 20 طالب و بفرض أن قيم العينة 6,7,8 . يعتبر الوسط الحسابي إحصاء وعادة يمثل بالرمز \bar{X} . قيمة المتغير العشوائي \bar{X} لهذا المثال هي \bar{x} ، وعلى ذلك :-

$$\bar{x} = \frac{6+7+8}{3} = 7.$$

في التطبيق، قيمة الإحصاء تستخدم في تقدير قيمة معلمة المجتمع . لمعرفة مدى جودة قيمة الإحصاء (التقدير) لابد من معرفة التوزيع الاحتمالي للإحصاء والذي يسمى التوزيع العيني sampling distribution . التوزيعات لبعض الإحصاءات المفيدة سوف نتاولها في الفصل السادس.

Frequency Distribution

(٢-٥) التوزيع التكراري

غالباً ما يكون التوزيع الاحتمالي لمتغير عشوائي غير معروف. تعتبر البيانات الإحصائية التي يجمعها الباحث بكميات كبيرة مفيدة جداً في دراسة سلوك المتغير العشوائي إذا تم عرضها بشكل مناسب. المعلومات الكثيرة يمكن الحصول عليها بتجميع البيانات في فئات classes وحساب عدد المشاهدات في كل فئة. مثل هذا التنظيم يسمى التوزيع التكراري . عندما نقوم بتجميع البيانات في فئات فإننا نحصل على أحسن صورة للمجتمع موضع الدراسة ولكننا في المقابل نفقد الكثير من التفاصيل عن المشاهدات في العينة. عدد المشاهدات في فئة خاصة يسمى تكرار الفئة class frequency ويرمز له بالرمز f . الجدول (١-٥) يمثل التوزيع التكراري لأطوال 22 نبات من نوع ما (المشاهدات معطاة لأقرب عدد صحيح). في هذا المثال لدينا 6 فئات وهم: 35-39 , 40-44 , 45-49 , 50-54 , 55-59 , 60-64 . يشار إلى أصغر وأكبر القيم التي تقع في فئة معطاة بمحدود هذه الفئة class limits . فعلى سبيل المثال للفئة 55-59 - أصغر رقم هو 55 ويمثل الحد الأدنى للفئة lower class limit وأكبر رقم هو 59 ويمثل الحد الأعلى للفئة upper class limit . وحيث أن البيانات الأصلية مسجلة لأقرب رقم

صحيح، فإن 4 مشاهدات تقع في الفئة 55-59 يمثلون كل المشاهدات في العينة التي قيمهم أكبر من أو يساوي 54.5 وأصغر من 59.5. الأرقام 54.5 و 59.5 تسمى الحدود الفعلية **class boundaries** للفئة 55-59. الرقم 54.5 يسمى الحد الأدنى الفعلي **lower class boundary** والرقم 59.5 يسمى الحد الأعلى الفعلي **upper class boundary**. أيضا الرقم 59.5 يسمى الحد الأدنى الفعلي للفئة التالية أي الفئة 60-64. ويلاحظ أنه بالرغم من أن الفئات لها حدود فعلية مشتركة إلا أنه من غير الممكن أن تقع مشاهدة واحدة على أحد هذه الحدود وذلك لأن الحدود الفعلية للفئات تحتوي على خانة عشرية أكبر من تلك الموجودة في البيانات نفسها.

جدول (١-٥)

حدود الفئة	35-39	40-44	45-49	50-54	55-59	60-64
التكرار	1	2	3	4	4	8

يعرف الفرق بين الحد الأعلى الفعلي والحد الأدنى الفعلي للفئة بطول الفئة **class width** ويساوي أيضا الفرق بين الحد الأعلى والحد الأدنى للفئة زائدا وحدة دقة، أي وحدة من الوحدات التي قربت إليها الأعداد في البيانات (في هذا المثال وحدة الدقة هي الواحد الصحيح لأننا قربنا البيانات لأقرب رقم صحيح). من الناحية العملية يفضل الحصول على فئات ذات أطوال متساوية ما أمكن. سوف نرمز لطول الفئة بالرمز Δ . أطوال الفئات في جدول (١-٥) متساوية وتساوي $\Delta = 5$.

منتصف الفئة **midpoint** تسمى مركز الفئة **class midpoint** أو **class mark** ونحصل عليها بجمع الحد الأدنى الفعلي والحد الأعلى الفعلي للفئة وقسمة المجموع على 2 وذلك تحت فرض أن جميع المشاهدات داخل الفئة تأخذ قيمة تتطابق مع مركز الفئة. مثال ذلك إفتراض أن 8 تكرارات في الفئة 60-64 تأخذ القيمة 62 والتي تمثل مركز هذه الفئة. أيضا يمكن الحصول على مركز الفئة بجمع الحد الأدنى والحد الأعلى للفئة وقسمة المجموع على 2. من جدول (١-٥) مراكز الفئات هم : 37, 42, 47, 52, 57, 62. يمثل جدول (١-٥) توزيع تكراري من النوع الذي نشاهده في التقارير المنشورة في الصحف. للأغراض الإحصائية يكون من الأفضل الحصول على توزيعات ذات تفصيلات أكثر، كما هو موضح في جدول (٢-٥) لنفس البيانات المعطاة في جدول (١-٥).

التكرار	مركز الفئة	الحدود الفعلية الفئة	حدود الفئة
1	37	34.5-39.5	35-39
2	42	39.5-44.5	40-44
3	47	44.5-49.5	45-49
4	52	49.5-54.5	50-54
4	57	54.5-59.5	55-59
8	62	59.5-64.5	60-64

مثال (١-٥) في عينة عشوائية حجمها $n=35$ من حيوانات التجارب تم تسجيل كمية مركب ما في الدم لكل حيوان كما في جدول (٣-٥) . المطلوب عرض هذه البيانات في توزيع تكراري .

جدول (٣-٥)

1.1	1.2	1.4	1.1	1.2	1.5	1.6
1.2	1.6	1.5	1.8	1.9	1.8	1.7
1.4	1.5	2.1	2.2	2.2	2.3	2.4
2.5	2.6	2.7	2.8	2.9	2.9	2.9
2.9	2.8	2.8	2.7	2.7	2.6	2.6

الحل . في البداية نقرر عدد الفئات والتي سوف توزع فيها البيانات . عادة يفضل أن تكون عدد الفئات من 5 إلى 20 . إذا زاد عدد الفئات عن عشرين خسر الباحث البساطة التي يكسبها عادة عند وضع البيانات في التوزيع التكراري ، وإذا قل عدد الفئات عن 5 فإن ذلك يؤدي إلى ضياع الكثير من التفاصيل الموجودة في البيانات . بفرض أننا قررنا أن يكون عدد الفئات 5 ، لحساب طول الفئة فإننا أولاً نحسب المدى وهو عبارة عن الفرق بين أكبر وأصغر مشاهدة في العينة وعلى ذلك يكون المدى $1.8 = 2.9 - 1.1$. لذا نقسم المدى على عدد الفئات المقترحة أي $\frac{1.8}{5} = 0.36$ ويقرب الكسر الناتج من خارج القسمة إلى أقرب رقم عشري (لأن البيانات الخام أصلاً مقاس لأقرب رقم عشري) . أي أن طول الفئة سوف يكون $\Delta = 0.4$. نحدد بداية الفئة الأولى (الحد الأدنى للفئة الأولى) والذي غالباً ما يكون أصغر رقم في البيانات وهو 1.1 ، وكذلك نحدد الحد الأدنى للفئة الثانية بإضافة طول الفئة إلى الحد الأدنى للفئة الأولى ، وهكذا لتعيين الحدود الدنيا لباقي الفئات . أما بالنسبة لتحديد الحد الأعلى للفئة الأولى فإنه يمكن تعيينه بإضافة طول الفئة إلى الحد الأدنى للفئة الأولى ثم نطرح من حاصل الجمع مقدار وحدة دقة من الوحدات التي قربت إليها الأعداد في المشاهدات ، أي 0.1 . وكذلك نحدد الحد الأعلى للفئة الثانية بإضافة طول الفئة إلى الحد الأعلى للفئة الأولى ، وهكذا لتعيين الحدود العليا لباقي الفئات ،

وذلك تحت شرط أن الفئات متساوية الأطوال. أخيرا نقوم بعدد المشاهدات التي تقع في كل فئة ويوضع الرقم في عمود التكرار، ولا بد أن يكون مجموع التكرارات مساويا لعدد المشاهدات في جدول (٣-٥) • يمثل جدول (٤-٥) التوزيع التكراري للبيانات المعطاة في جدول (٣-٥). التكرار النسبي لكل فئة يمكن الحصول عليه بقسمة تكرار الفئة على مجموع التكرارات. الجدول الذي يحتوي على التكرارات النسبية يسمى التوزيع النسبي **the relative frequency** • التوزيع النسبي للبيانات المعطاة في جدول (٤-٥) موضح في جدول (٥-٥) • وضرب كل تكرار نسبي في 100 نحصل على التكرار النسوي **percentage distribution** •

جدول (٤-٥)

حدود الفئة	الحدود الفعلية الفئة	مركز الفئة	التكرار
1.1-1.4	1.05-1.45	1.25	7
1.5-1.8	1.45-1.85	1.65	8
1.9-2.2	1.85-2.25	2.05	4
2.3-2.6	2.25-2.65	2.45	6
2.7-3.0	2.65-3.05	2.85	10
المجموع			35

جدول (٥-٥)

حدود الفئة	1.1-1.4	1.5-1.8	1.9-2.2	2.3-2.6	2.7-3.0	المجموع
التكرار النسبي	0.2000	0.2286	0.1143	0.1714	0.2857	1

في بعض الأحيان يكون الإهتمام ليس فقط بعدد المشاهدات في فئة معطاة ولكن في عدد المشاهدات الذي يقع فوق أو تحت قيمة معينة. على سبيل المثال في جدول (٤-٥) عدد الحيوانات التي كمية المركب في دمها 2.25 أو أقل هو $7+8+4=19$ ، والذي يمثل التكرار المتجمع **cumulative frequency** للفئة الرابعة. جدول (٦-٥) يوضح التكرارات المتجمعة والتي تم حسابها من جدول (٤-٥) ، ويسمى التوزيع التكراري المتجمع **cumulative frequency distribution** .

أما إذا استخدما التكرارات النسبية بدلا من التكرارات فإننا نحصل على التكرار المتجمع النسبي **relative cumulative frequency** . وإذا استخدما التكرارات النسبية بدلا من التكرارات فإننا نحصل على التوزيع التكراري المتجمع النسوي • كما يمكن الحصول

التوزيع التكراري المتجمع المتوي من جدول التكرار المتجمع وذلك بقسمة التكرار المتجمع لكل فئة على مجموع التكرارات وضرب الناتج في 100% . يوضح جدول (٧-٥) التوزيع التكراري المتجمع المتوي للبيانات في جدول (٦-٥) .

جدول (٦-٥)

الحدود العليا للفئات	التكرار المتجمع
أقل من 1.05	0
أقل من 1.45	7
أقل من 1.85	15
أقل من 2.25	19
أقل من 2.65	25
أقل من 3.05	35

جدول (٧-٥)

الحدود العليا للفئات	التكرار المتجمع %
أقل من 1.05	00.00
أقل من 1.45	20.00
أقل من 1.85	42.86
أقل من 2.25	54.29
أقل من 2.65	71.43
أقل من 3.05	100.00

مثال (٢-٥) البيانات في جدول (٨-٥) تمثل الدخل اليومي الصافي - لأقرب جنيه - في محل تجاري في مدة 40 يوم. كون جدول تكراري لهذه البيانات على أن يحصر الجدول على مركز - الفئة - الحدود الفعلية للفئة - التكرار المتوي - التكرار المتجمع .

جدول (٨-٥)

118	124	128	134	135	138	140	142
125	130	136	138	141	143	145	144
144	146	147	150	152	154	155	146
146	147	155	168	157	163	160	163
146	168	170	175	181	181	175	168

الحل . أولا نحسب المدى = أكبر قيمة - أصغر قيمة أي $181-118=63$ وحيث أن عدد الفئات المقترحة 13 فإن طول الفئة هو $\frac{63}{13} = 4.846$ أي تقريبا 5 لأن البيانات مقاسة لأقرب عسدد صحيح. الحد الأدنى للفئة الأولى هو 118 والذي يمثل أصغر قيمة في البيانات. جدول (٩-٥)
يحتوى على المطلوب.

جدول (٩-٥)

الحدود الفعلية	مركز الفئة	التكرار	التكرار المتوي	التكرار المتجمع
118-122	120	1	2.5	1
123-127	125	2	5.0	3
128-132	130	2	5.0	5
133-137	135	3	7.5	8
138-142	140	5	12.5	13
143-147	145	10	25.0	23
148-152	150	2	5.0	25
153-157	155	4	10.0	29
158-162	160	1	2.5	30
163-167	165	2	5.0	32
168-172	170	4	10.0	36
173-177	175	2	5.0	38
178-182	180	2	5.0	40
المجموع		40	100	

مثال (٣-٥) أخذت عينة من مزرعة دواجن وكانت أوزان الدجاج مقاسة لأقرب مائة جرام كما في جدول (١٠-٥). كون جدول تكراري لهذه البيانات على أن يحتوى الجدول على مركز الفئة - الحدود الفعلية للفئة - التكرار المتوي - التكرار المتجمع .
الحل . أولا نحسب المدى = أكبر قيمة - أصغر قيمة أي $1400-600=800$ وحيث أن عدد الفئات المقترحة 5 فإن طول الفئة هو $\frac{800}{5} = 160$ أي تقريبا 200 لأن البيانات مقاس لأقرب 100 جرام. الحد الأدنى للفئة الأولى هو 600 والذي يمثل أصغر قيمة في البيانات. جدول (١١-٥)
يحتوى على المطلوب.

جدول (١٠-٥)

600	600	900	1000	800	800	900
700	800	800	900	700	700	700
600	900	700	1100	1200	1200	1300
1300	1300	1400	1000	1200	1200	1400
1300	1200	1300	1300	1400	1400	1000

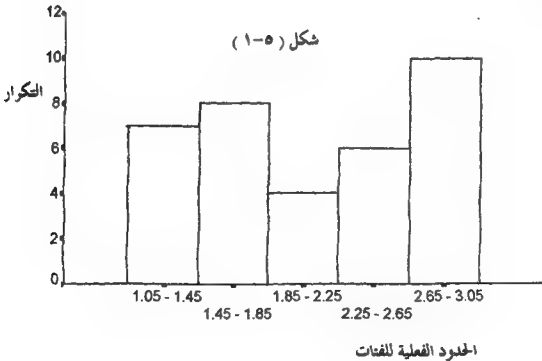
جدول (١١-٥)

الحدود الفعلية	الحدود الفعلية	مركز الفئة	التكرار	التكرار المتوي	التكرار المتجمع
600-700	550-750	650	8	22.86	8
800-900	750-950	850	8	22.86	16
1000-1100	950-1150	1050	4	11.43	20
1200-1300	1150-1350	1250	11	31.42	31
1400-1500	1350-1550	1450	4	11.43	35
المجموع			35	100	

Graphic Representation

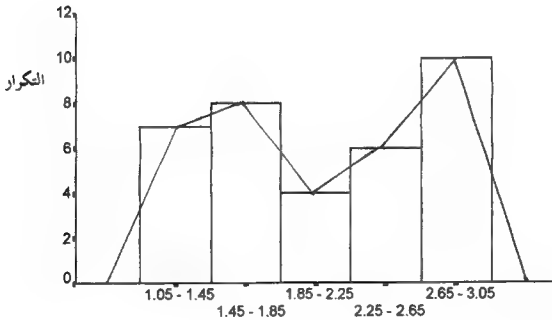
(٣-٥) التمثيل البياني

المعلومات التي نحصل عليها من التوزيع التكراري في شكل جداولي تصبح أسهل في الفهم إذا تم عرضها بيانياً . من أكثر الأشكال البيانية الواسعة الانتشار في تمثيل البيانات الرقمية ما يعرف بالمدرج التكراري histogram والذي يناسب البيانات المتصلة. ونحصل عليه بتمثيل تكرار كل فئة من فئات التوزيع بمستطيل قاعدته الحدود الفعلية لتلك الفئة وارتفاعه يساوي تكرار الفئة. ويتم ذلك برسم محورين أحدهما أفقي والآخر رأسي ونرصد على المحور الأفقي الحدود الفعلية لكل فئة من فئات التوزيع التكراري ونقيم على كل فئة مستطيل ارتفاع يساوي تكرار تلك الفئة. المدرج التكراري للتوزيع التكراري المعطى في جدول (٤-٥) موضح في شكل (١-٥) .



في بعض المشاكل يكون من الأفضل وضع التكرار النسبي أو النسبي على المحور الرأسي. المدرج التكراري في هذه الحالة يسمى المدرج التكراري النسبي **relative frequency histogram** ، **percentage frequency histogram** أو المدرج التكراري النسبي ، ويكون لهما نفس شكل المدرج التكراري، كما أن مجموع مساحات الأعمدة للمدرج التكراري النسبي تساوي الواحد الصحيح.

الطريقة الثانية المفيدة لتمثيل البيانات الرقمية بياناتها هو استخدام المضلع التكراري **frequency polygon** والذي نحصل عليه بتصنيف الأضلاع العلوية للمستطيلات في المدرج التكراري ثم نوصل هذه النقاط بعضها البعض. ولكي نغلق الخط المنكسر الذي حصلنا عليه نحدد على محور الأفقي مركز الفئة السابقة للفئة الأولى ومركز الفئة اللاحقة للفئة الأخيرة ونغلق المضلع. يوضح شكل (٢-٥) المضلع التكراري للتوزيع التكراري في جدول (٤-٥). ويوضح من الشكل أن مجموع مساحات المستطيلات يساوي المساحة تحت المضلع التكراري، ويحدث هذا فقط في حالة الفئات المتساوية .



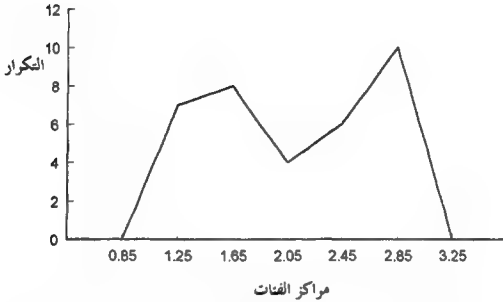
الحدود الفعلية للفئات
شكل (٢-٥)

وهناك طريقة أخرى لرسم المضلع التكراري وذلك برسم محورين أحدهما أفقي والآخر رأسي. يمثل المحور الأفقي مراكز الفئات ويمثل المحور الرأسي التكرارات. نعتبر مركز كل فئة إحداثياً أفقياً لنقطة ونعتبر تكرار هذه الفئة الإحداثي الرأسي لتلك النقطة. ولكي نغلق الخط

المتكسر الذي حصلنا عليه نحدد على محور الألفي مركز الفئة السابقة للفئة الأولى ومركز الفئة اللاحقة للفئة الأخيرة ونغلق المضلع. يوضح شكل (٣-٥) المضلع التكراري للتوزيع التكراري في جدول (٤-٥) وذلك بتحديد النقاط الآتية على الرسم :-

(2.85,10), (2.45,6), (2.05,4), (1.65,8), (1.25,7) والنقطتان الإضافيتان هما (0.85,0), (3.25,0). ثم وصل هذه النقاط بعضها ببعض .

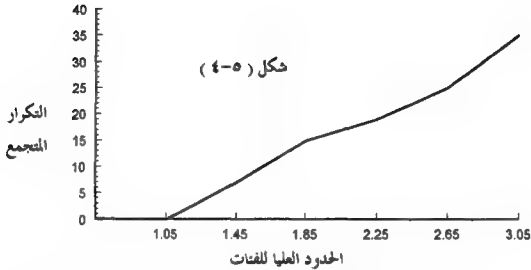
شكل (٣-٥)



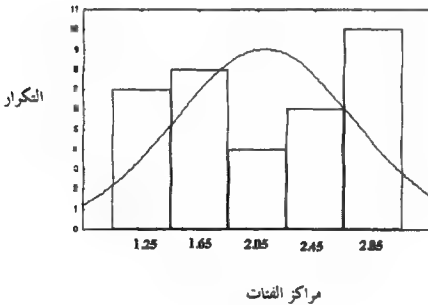
عند الرغبة في مقارنة فئتين من البيانات مختلفتين في عدد مفرداتهما فإنه يمكن تمثيل المضلعين التكراريين على نفس الرسم. في هذه الحالة لابد من استخدام التكرارات النسبية أو المتوية .

أما بالنسبة للتوزيعات التكرارية المتجمعة فهناك ما يسمى المضلع التكراري المتجمع cumulative frequency polygon ونحصل عليه برسم محورين أحدهما أفقي والآخر رأسي وكل نقطة على الرسم إحداثياتها الحد الأعلى الفعلي للفئة والتكرار المتجمع وتوصل النقاط نحصل على المضلع التكراري المتجمع. شكل (٤-٥) يوضح المضلع التكراري المتجمع للتوزيع التكراري المتجمع في جدول (٥ - ٦).

يمكن تمثيل التوزيع التكراري المتجمع النسبي والتوزيع التكراري المتجمع المتوي بيانياً، بنفس الطريقة التي مثلنا بها التوزيع التكراري المتجمع بيانياً، وذلك باستخدام التكرارات المتجمعة النسبية والتكرارات المتجمعة المتوية .



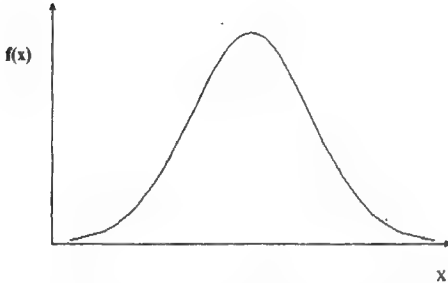
هناك طريقة أخرى لتمثيل التوزيعات التكرارية بيانياً وذلك باستخدام المنحنى التكراري frequency curve . ونحصل عليه برسم المضلع التكراري ونمهد الخطوط المنكسرة التي تصل بين هذه النقاط. وقد يكون التمهيد باليد أو بطرق رياضية ولا يشترط أن يمر المنحنى بجميع رؤوس المضلع التكراري. يبين شكل (٥-٥) المدرج التكراري والمنحنى التكراري معاً للتوزيع التكراري في جدول (٤-٥) ، عموماً كلما ضاقت أطوال الفئات وزاد عدد المشاهدات لمسإن المضلع التكراري يؤول إلى المنحنى التكراري.



شكل (٥-٥)

عند الرغبة في تقدير التوزيع الاحتمالي لمغير عشوائي متصل X نقوم بمجهد المنحنى التكراري النسبي . شكل المنحنى يساعدنا في اقتراح شكل $f(x)$ للمتغير العشوائي المتصل موضع الدراسة. على الرغم من أننا تمكنا من الحصول على تقدير للدالة $f(x)$ بيانيا فلا نزال نهمل الصيغة الرياضية أو المعادلة الخاصة بالدالة $f(x)$ وبالتالي لا نستطيع حساب تقديرات للاحتمالات. كثير من التوزيعات المتصلة يمكن تمثيلها بيانيا بمنحنى على شكل الناقوس **bell** كما في شكل (٦-٥) . الصيغة أو المعادلة الخاصة بالدالة $f(x)$ معروفة (دالة كثافة الاحتمال) وتعتمد على معلمتين μ , σ . بمجرد الحصول على تقدير لكل منهما من البيانات يمكن كتابة المعادلة المقدرة ثم استخدام الجداول المناسبة لحساب أي مساحة تحت المنحنى .

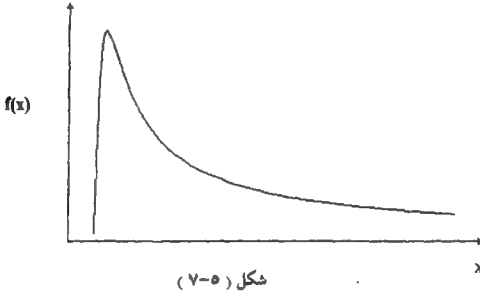
شكل (٦-٥)



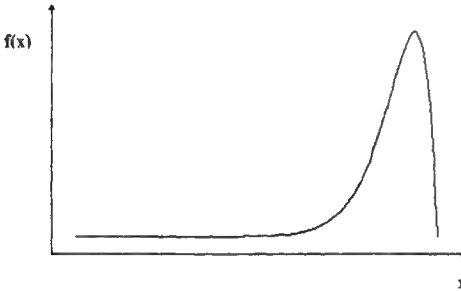
عادة تأخذ المنحنيات أشكال كثيرة ويمكن اشتقاق معادلتها بتقدير معالها المجهولة. مشاكل التقدير صعبة جدا وتنتمي إلى فرع الإحصاء الرياضي.

يقال للتوزيع أنه متماثل **symmetrical** ، كما في شكل (٦-٥) إذا أمكننا إقامة عمود على محور الأفقي بحيث يقسم هذا العمود التوزيع إلى قسمين ينطبقان على بعضهما تمام الانطباق. النقطة التي تمكنا من إقامة العمود تسمى نقطة التماثل. أما التوزيعات التي يكون عدم التماثل واضحا فتسمى توزيعات ملتوية **skewed** . يكون التوزيع ملتويا إلى اليمين أو موجب الالتواء **positive skewed** إذا كان معدل التناقص في المنحنى أسرع جهة اليمين منه جهة اليسار بحيث يكون الجانب الأيمن من المنحنى أطول من الجانب الأيسر كما في شكل (٦-٥)

• بينما يكون التوزيع ملتويا إلى اليسار وسالب الالتواء **negative skewed** إذا كان معدل التناقص في المنحنى أسرع جهة اليسار منه جهة اليمين بحيث يكون الجانب الأيسر من المنحنى أطول من الجانب الأيمن كما في شكل (٨-٥) .



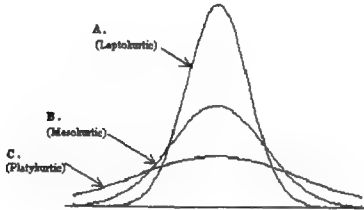
شكل (٨-٥)



شكل (٩-٥)

عند مقارنة المنحنيات وحيدة القمة قد نجد أنها تختلف من حيث شكل القمة. فقد تكون قمة إحداها أكثر تدبها أو تفرطحاً من بعض القمم الأخرى . ففي شكل (٩-٥) ثلاث منحنيات مختلفة في كمية التفلطح . المنحنى ، A ، ذو القمة المدببة **leptokurtic** يمثل توزيع بقيم تتركز بشدة حول نقطة الوسط **midpoint** . المنحنى الثاني ، B ، وهو المعتدل

mesokurtic يكون متوسط التفلطح ويمثل توزيع بقمم تتركز بدرجة أقل حول نقطة الوسط عن المنحني المدب • وأخيرا المنحني ، C ، الفلطح platykurtic والذي يكون منبسطا وتنخفض قمته عن قمة المنحني المعتدل • وهذا يدل على أن قيمة تقع حول نقطة الوسط في مدى غير ضيق •



شكل (٩-٥)

Measures of Central Tendency

(٤ - ٥) مقاييس الرعة المركزية

في بعض الأحيان التمثيل البياني وحده لا يمد الباحث بكل المعلومات التي يحتاج إليها من فئة المشاهدات تحت الدراسة، فالبيانات لابد أن توصف وتحلل. واحد من الطرق لوصف فئة من المشاهدات، سواء عينة أو مجتمع، هو استخدام المتوسطات averages (مقاييس الرعة المركزية) ، فالمتوسط هو القيمة التي تتركز حولها معظم المشاهدات . في هذا البند سوف نستعرض أربعة مقاييس للرعة المركزية .

Arithmetic Mean

(١ - ٤ - ٥) الوسط الحسابي

يعتبر الوسط الحسابي من أفضل مقاييس الرعة المركزية لتمثيل مركز فئة من المشاهدات. تعريف : إذا كانت الفئة من المشاهدات x_1, x_2, \dots, x_N ، ليس من الضروري أن تكون كلها مختلفة ، تمثل مشاهدات مجتمع محدود من الحجم N ، فإن الوسط الحسابي للمجتمع يمكن حساب من الصيغة التالية :-

$$\mu = \frac{\sum_{i=1}^N x_i}{N}.$$

مثال (٤-٥) أوجد الوسط الحسابي لمتجمع مشاهداته هي :-

$$8, 10, 13, 9, 7, 11, 10, 12, 10, 9, 11.$$

الحل.

$$\begin{aligned}\mu &= \frac{\sum_{i=1}^N x_i}{N} = \frac{8+10+13+9+7+11+10+12+10+9+11}{11} \\ &= \frac{110}{11} = 10.\end{aligned}$$

تعريف : إذا كانت الفئة من المشاهدات x_1, x_2, \dots, x_n ، ليس من الضروري أن تكون كلها مختلفة ، تمثل مشاهدات عينة من الحجم n ، فإن الوسط الحسابي للعينة يمكن حسابه من الصيغة التالية :-

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}.$$

مثال (٥-٥) أوجد الوسط الحسابي لعينة مشاهداتها هي :-

$$6, 7, 7, 8.$$

الحل .

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^n x_i}{n} = \frac{6+7+7+8}{4} \\ &= \frac{28}{4} = 7.\end{aligned}$$

عند وضع البيانات في توزيع تكراري فإننا نفقد الهوية لأي مشاهدة في العينة.

المعلومات التي تبقى هي عدد المشاهدات التي تقع في كل فئة. لحساب الوسط الحسابي من توزيع تكراري نفترض أن كل المشاهدات داخل فئة معطاة تقع عند مركز الفئة.

تعريف : إذا كانت x_1, x_2, \dots, x_k هي مراكز الفئات لتوزيع تكراري مع تكراراتها المقابلة

f_1, f_2, \dots, f_k (حيث k تمثل عدد الفئات) فإن الوسط الحسابي يحسب من الصيغة التالية :

$$\bar{x} = \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i}$$

مثال (٦-٥) أوجد الوسط الحسابي للبيانات المعطاة في جدول (١٢-٥) والتي تمثل التوزيع التكراري لأطوال عينة من الأسماك.

جدول (١٢-٥)

حدود الفئة	التكرار	مركز الفئة	
	f_i	x_i	$f_i x_i$
30-39	11	34.5	379.5
40-49	12	44.5	534.0
50-59	16	54.5	872.0
60-69	23	64.5	1483.5
70-79	17	74.5	1266.5
80-89	11	84.5	929.5
90-99	10	94.5	945.0
المجموع	100		6410

الحل . بما أن $k = 7$, $\sum_{i=1}^k f_i = 100$, $\sum_{i=1}^k f_i x_i = 6410$ وعلى ذلك الوسط الحسابي هو :

$$\bar{x} = \frac{\sum_{i=1}^k f_i x_i}{\sum_{i=1}^k f_i} = \frac{6410}{100} = 64.1.$$

في بعض الأحيان يكون من المناسب إضافة (أو طرح) ثابت إلى القيم الأصلية ثم حساب الوسط الحسابي. السؤال الآن كيف تكون العلاقة بين الوسط الحسابي الجديد والوسط الحسابي لفئة المشاهدات الأصلية ؟ من نظرية (١-٤) بوضع $a = 1$ فإن $E(X + b) = E(X) + b$ ، وعلى ذلك فإن إضافة (أو طرح) ثابت إلى كل القيم سوف يغير الوسط الحسابي بنفس المقدار . على سبيل المثال لإيجاد الوسط الحسابي للقيم 0,2,5,8,10 والتي وسطها الحسابي 5. وعلى ذلك فإن الوسط الحسابي للقيم الأصلية $5 - 4 = 1$. أيضا من نظرية (١-٤) و نتيجة (٢) فإن $E(aX) = aE(X)$. وعلى ذلك بضرب (أو قسمة) ثابت إلى فئة من البيانات، فإن المشاهدات الأصلية سوف يكون وسطها الحسابي عبارة عن

الوسط الحسابي للملاحظات الجديدة مقسوما على (مضروباً في) الثابت . على سبيل المثال
الوسط الحسابي للقيم 5,10,15 هو 10، وعلى ذلك بعد قسمة كل الملاحظات على 5، فإن
الملاحظات الجديدة هي 1,2,3 والتي وسطها 2 . وعلى ذلك الوسط الحسابي للقيم الأصلية هو
 $10 = (5)(2)$.

ومن مميزات الوسط الحسابي أنه مألوف وسهل الفهم كما أنه معروف لأي فئة من البيانات
وقيمته وحيدة . أيضا كل قيمة في فئة الملاحظات تدخل في حسابه .

أما عيوبه فهي تأثره بالقيم الشاذة ولذلك لا ينصح باستخدامه للبيانات التي منحناها
شديد الالتواء. أيضا لا يمكن تقديره من التوزيعات التكرارية التي تحوى على فئات مفتوحة .
وأخيرا لا يمكن حسابه بالرسم .

ومن الخصائص المميزة للوسط الحسابي أن مجموع انحرافات القيم (الانحراف هو بعد أي

$$\sum_{i=1}^n (x_i - \bar{x}) = 0 \text{ أي أن } \bar{x} \text{ هو الوسط الحسابي يساوى صفراً، أى أن } \sum_{i=1}^n (x_i - \bar{x}) = 0$$

، أما في حالة التوزيعات التكرارية فإن $\sum_{i=1}^k (x_i - \bar{x}) \cdot f_i = 0$. أيضا مجموع مربعات الانحرافات
القيم حول الوسط الحسابي $\sum_{i=1}^n (x_i - \bar{x})^2$ أقل ما يمكن، أى أقل من مجموع مربعات انحرافات
القيم عن أى قيمة أخرى غير الوسط الحسابي .

عند حساب الوسط الحسابي يفترض أن كل قيمة لها نفس الأهمية، مثل هذا القرض قد
يكون خاطئاً . في الحقيقة، إذا كانت القيم ليس لها نفس الأهمية يكون من الأفضل حساب
الوسط الحسابي المرجح. فإذا كانت x_1, x_2, \dots, x_n تمثل قيم المتغير X ، وكانت
الأوزان المناظرة لها فإن الوسط الحسابي المرجح يكون :-

$$\bar{x}_w = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$$

مثال (٥-٧) تدفع شركة أجراً قدره 25 جنيهاً في الساعة لعمالها غير مهرة وعددهم 20
وتدفع 30 جنيهاً في الساعة للعمال شبه مهرة وعددهم 10 وتدفع 40 جنيهاً في الساعة للعمال
المهرة وعددهم 5 . ما هو الوسط الحسابي المرجح للأجر في الساعة التي تدفعه الشركة ؟
الحل .

$$\bar{x}_w = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} = \frac{(25)(20) + (30)(10) + (40)(5)}{20 + 10 + 5} = \frac{1000}{35} = 28.57.$$

يعاب على الوسط الحسابي المرجح هو عدم وجود قاعدة لتحديد الأوزان وتخضع للتقدير الشخصي. يعتبر الوسط الحسابي حالة خاصة من الوسط الحسابي المرجح إذا كانت الأوزان متساوية .

إذا كانت لدينا k من المجموعات، وكانت أحجام عيناتها n_1, n_2, \dots, n_k وأوساطها الحسابية على التوالي $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ فإن الوسط الحسابي المرجح للمجموعة التي حجمها $(n_1 + n_2 + \dots + n_k)$ الناتجة من دمج k من المجموعات هو :-

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k n_i \bar{x}_i}{\sum_{i=1}^k n_i}.$$

ويسمى الوسط الحسابي المرجح للأوساط الحسابية .

مثال (٨-٥) إذا كان الوسط الحسابي لأطوال 20 حيوان من نوع ما هو $\bar{x}_1 = 10$ وكان الوسط الحسابي لأطوال 30 حيوان من مجموعة أخرى من نفس النوع هو $\bar{x}_2 = 15$ أوجد الوسط الحسابي المرجح للأوساط الحسابية .
الحل .

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k n_i x_i}{\sum_{i=1}^k n_i} = \frac{(20)(10) + (30)(15)}{20 + 30} = \frac{650}{50} = 13.$$

Median (٥-٤-٢) الوسيط

يعتبر الوسيط هو المقياس الأفضل بعد الوسط الحسابي . سوف نرمز لوسيط المجتمع بالرمز \tilde{x} ووسيط العينة بالرمز \bar{x} . الوسيط لفئة من المشاهدات مرتبة تصاعدياً (أو تنازلياً) هو العدد الأوسط منها إذا كان عددها فردياً وهو الوسط الحسابي للعديد الأوسطين إذا كان عددها زوجياً .

تعريف : إذا كانت x_1, x_2, \dots, x_n تمثل فئة من المشاهدات المرتبة تصاعديا (أو تنازليا)

فإن الوسيط لهذه الفئة هو العدد $x_{(\frac{n+1}{2})}$ إذا كان n فرديا وهو العدد $\frac{1}{2}[x_{(\frac{n}{2})} + x_{(\frac{n+2}{2})}]$ إذا كان n زوجيا .

إذا كانت n زوجيا .

مثال (٩-٥) أوجد الوسيط للمجتمع الذي مشاهداته 10,9,8,6,7

الحل . بترتيب المشاهدات تصاعديا أي 6,7,8,9,10 فإن وسيط المجتمع يكون $\tilde{x} = 8$.

مثال (١٠-٥) أوجد الوسيط للعينة التي مشاهداتها هي 10,9,6,1,2,7

الحل . بترتيب البيانات تصاعديا أي 1,2,6,7,9,10 فإن وسيط العينة يكون :

$$\bar{x} = \frac{6+7}{2} = \frac{13}{2} = 6.50$$

ومن الخصائص المميزة للوسيط أن مجموع القيم المطلقة لانحرافات القيم عن الوسيط أقل

ما يمكن. أي أن ، $|x_i - \bar{x}|$ أقل ما يمكن . القيمة المطلقة للرقم هي القيمة بدون إشارة ، وعلى ذلك فإن $|-5|=5$.

من مميزات الوسيط أنه سهل الفهم ولا يتأثر بالقيم الشاذة. كما يمكن استخدامه في

التوزيعات التي تحتوي على فئات مفتوحة .

ومن عيوب الوسيط أن كل المشاهدات لا تدخل في حسابه . كما أنه مثل الوسيط

الحسابي ، في بعض الأحيان ، يكون قيمة صناعية artificial ، بمعنى عدم وجود قيمة في فئة

المشاهدات ، في الحقيقة ، تمثل الوسيط.

تعريف : لأي توزيع تكراري فإن الفئة التي تحتوي على الوسيط تسمى الفئة

الوسيط median class .

مثال (١١-٥) أوجد الوسيط للتوزيع التكراري المعطى في جدول (١٢-٥) .

الحل . لحساب الوسيط نفترض أن البيانات تتوزع بانتظام داخل فئة الوسيط . ومن جدول (

١٢-٥) نحصل على جدول (١٣-٥) . لحساب الفئة الوسيطة نحسب موقع الوسيط وهو من

جدول (١٣-٥) يساوي $\frac{n}{2} = \frac{100}{2} = 50$ (حيث $\sum_{i=1}^k f_i = n$) سواء كان مجموع

التكرارات فرديا أو زوجيا . من جدول (١٣-٥) وفي عمود التكرار المتجمع نجد أن 62

مشاهدة قيمهم أقل من 69.5 . أيضا من عمود التكرار المتجمع نجد أن 39 مشاهدة قيمهم أقل

من 59.5 . أي أن الوسيط لابد أن يقع في الفئة الرابعة ، أي أن فئة الوسيط هي 59.5-59.5

69.5 • وما أن الوسيط هو القيمة التي ترتيبها الـ 11 = 39-50 داخل الفئة الوسيطة حيث 50 تمثل ترتيب الوسيط و 39 تمثل التكرار المتجمع السابق لفئة الوسيط، وتحت فـرض أن 23 مشاهدة تمثل تكرار الفئة الوسيطة موزعة بانتظام على الفئة الوسيطة التي طولها 10 وعلى ذلك تكون المسافة من بداية الفئة الوسيطة وموقع الوسيط هي 4.783 = $10 \cdot \frac{11}{23}$ •

جدول (٥ - ١٣)

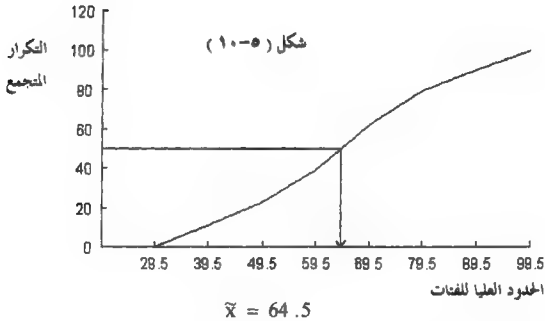
التكرار المتجمع	الحدود العليا للفئات	التكرار f_i	الحدود الفعلية
11	أقل من 39.5	11	29.5-39.5
23	أقل من 49.5	12	39.5-49.5
39	أقل من 59.5	16	49.5-59.5
62	أقل من 69.5	23	59.5-69.5
79	أقل من 79.5	17	69.5-79.5
90	أقل من 89.5	11	79.5-89.5
100	أقل من 99.5	10	89.5-99.5
		100	المجموع

أي أن الوسيط هو $59.5 + 4.783 = \bar{x}$ أي $\bar{x} = 64.283$ ، وعلى ذلك يمكن القول أن 50% من الأسماك أطولها تكون أقل من 64.283 ، ولأن الكثيرون يفضلون حساب الوسيط من صيغة فسوف نقدم لهم الصيغة التالية لحساب الوسيط :-

$$\bar{x} = L + \left(\frac{\frac{n}{2} - F}{f_m} \right) \Delta$$

حيث : L = الحد الأدنى الفعلي لفئة الوسيط ، F = التكرار المتجمع السابق لفئة الوسيط ، Δ = طول فئة الوسيط ، f_m = تكرار فئة الوسيط •

يمكن إيجاد الوسيط بالرسم من المصنع التكراري المتجمع. فعلى سبيل المثال للتوزيع التكراري لجدول (٥-١٣) نرسم المصنع التكراري المتجمع كما في شكل (٥-١٠) ويتم تحديد موقع الوسيط على المحور الرأسي ثم نسقط عمود من نقطة موقع الوسيط على المصنع التكراري المتجمع وعند التقائه بالمصنع نسقط عمود على المحور الأفقي فتكون هي قيمة الوسيط. من شكل (٥-١٠) فإن الوسيط يكون $\bar{x} = 64.5$.



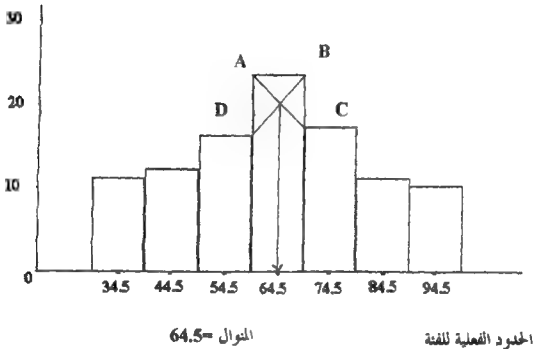
(٥-٤-٣) Mode المنوال

يعرف المنوال بأنه القيمة الأكثر شيوعاً أو التي تتكرر أكثر من غيرها. في بعض الأحيان لا يوجد منوال لفئة من المشاهدات حيث لا تتكرر القيم أكثر من مرة، وإذا وجد قد لا يكون وحيداً. على سبيل المثال المنوال للملاحظات 3,5,5,5,5,5,7,7,9 هو 5. مثال (٥-١٢) أوجد المنوال للملاحظات 2,4,5,6,6,6,6,7,7,7,7,8,9. الحل: التوزيع في هذه الحالة يكون ثنائي المنوال حيث المنوال هو 6,7.

يعتبر المنوال أقل مقياس الرتبة المركزية استخداماً. للفئات الصغيرة من البيانات لا يكون له فائدة، فقط يكون له معنى إذا كان حجم البيانات كبيراً. ومن مميزاته أنه لا يحتاج إلى عمليات حسابية، كما يمكن استخدامه للبيانات الوصفية. ويمكن حساب المنوال بالرسم من المدرج التكراري. فعلى سبيل المثال للتوزيع التكراري لجدول (٥-١٣) نرسم المدرج التكراري كما في شكل (٥-١١). نصل الرأس الأيمن العلوي للمستطيل الذي يمثل أكبر

تكرار بالرأس الأيمن للمستطيل السابق له (BD كما في الشكل) • أيضا نصل الرأس الأيسر العلوي لأطول مستطيل بالرأس الأيسر العلوي للمستطيل التالي له (AC) فيقاطع المستقيمان في نقطة، نسقط عمود من هذه النقطة على محور الألفي فيكون هو النوال • من شكل (٥-١١)

شكل (١١-٥)



في حالة البيانات التكرارية فإن النوال لدالة كثافة الاحتمال $f(x)$ هو قيمة x التي عندها يأخذ المنحنى أعلى قيمة • وعلى ذلك يمكن اعتبار مركز الفئة التي يقابلها أعلى تكرار هو تقدير للتوسط • وعلى ذلك للتوزيع التكراري في جدول (٥-١٢) فإن النوال التقريبي هو 64.5 •

The geometric Mean

(٥-٤-٤) الوسط الهندسي

في الحقيقة ، فإن الوسط الهندسي له استخدامات خاصة في المشاكل الاقتصادية وفي المجال السكاني •

تعريف : إذا كان لدينا الفئة من المشاهدات x_1, x_2, \dots, x_n ، فإن الوسط الهندسي يمكن حسابه من الصيغة التالية :-

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

ولتسهيل حساب الوسط الهندسي تستخدم الصيغة التالية إذا كان $n > 2$:-

$$\log G = \frac{\sum_{i=1}^n \log x_i}{n}$$

مثال (١٣-٥) حصل مستثمر على عائد من رأسماله المستثمر قدره 2% للسنة الأولى، 3% للسنة الثانية، 5% للسنة الثالثة أوجد الوسط الهندسي للعائد السنوي.
الحل •

$$\log G = \frac{\sum_{i=1}^n \log x_i}{n}$$

$$\begin{aligned} \log G &= \frac{\sum_{i=1}^n \log x_i}{n} = \frac{1}{3} (\log 2 + \log 3 + \log 5) \\ &= \frac{1}{3} (0.3010 + 0.4771 + 0.6990) \\ &= \frac{1.4771}{3} = 0.49237. \end{aligned}$$

وعلى ذلك فإن $G = 3.1072$ •

دائما يكون الوسط الهندسي أصغر من الوسط الحسابي • كما أن الوسط الهندسي لا يمكن حسابه إذا كانت إحدى القيم مساوية للصفر أو رقم سالب •

يمكن حساب الوسط الهندسي من جداول تكرارية من التعريف التالي •

تعريف : إذا كانت x_1, x_2, \dots, x_k تمثل مراكز الفئات لتوزيع تكراري مع تكرارها المقابل f_1, f_2, \dots, f_k (حيث k تمثل عدد الفئات) فإن الوسط الهندسي يحسب من الصيغة التالية :

$$\bar{x} = \frac{\sum_{i=1}^k f_i \log x_i}{\sum_{i=1}^k f_i}.$$

مثال (١٤-٥) البيانات في جدول (١٤-٥) تمثل التوزيع التكراري لأطوال عينة من حمسين نبات من نوع ما والمطلوب إيجاد الوسط الهندسي •

الحل • من جدول (١٤-٥) فإن :-

$$\bar{x} = \frac{\sum_{i=1}^k f_i \log x_i}{\sum_{i=1}^k f_i} = \frac{55.1799}{50} = 1.103598.$$

وعلى ذلك فإن الوسط الهندسي هو $G = 12.69399$.

جدول (١٤-٥)

حدود الفئة	التكرار f_i	مركز الفئة x_i	$\log x_i$	$f_i \log x_i$
6-8	6	7	0.8451	5.0706
9-11	10	10	1.0000	10.0000
12-14	15	13	1.1139	16.7085
15-17	12	16	1.2041	14.4492
18-20	7	19	1.2788	8.9516
المجموع	50			55.1799

(٥-٥) الربعات والثنيات والعشريرات

Quartiles, Percentiles, Deciles

كما ذكرنا سابقا، إذا رتبنا فئة من المشاهدات حسب قيمها تصاعديا فإن القيمة التي تكون في المنتصف والتي تقسم البيانات إلى قسمين متساويين في العدد هي الوسيط . وبتميم الفكرة وتقسيم البيانات إلى أربعة أجزاء متساوية (بعد ترتيب المشاهدات تصاعديا) فإن نقاط التقسيم يرمز لها بالرموز Q_1, Q_2, Q_3 حيث Q_1 يسمى الربع الأول first quartile (الربع الأدنى lower quartile) و Q_2 يسمى الربع الثاني second quartile (نفسه الوسيط) و Q_3 يسمى الربع الثالث third quartile (الربع الأعلى upper quartile) فالربع الأول هو القيمة Q_1 التي يسبقها ربع البيانات ويليهما ثلاثة أرباع البيانات . والربع الثاني (وهو أيضا الوسيط) هو القيمة Q_2 التي يسبقها نصف البيانات ويليهما نصف البيانات . وفي النهاية الربع الثالث وهو القيمة Q_3 التي يسبقها ثلاثة أرباع البيانات ويليهما ربع البيانات . عند استخدام فئة من المشاهدات فإن الربعات الثلاثة يتم حاسمها بتعين موقعها أولا فموقع الربع الأول هو $\frac{n+2}{4}$ والربع الثاني موقعه هو $\frac{n+1}{2}$ والربع الثالث موقعه هو $\frac{3n+2}{4}$.

للملاحظات في شكل (١٢-٥) موقع الربيع الأول هو $\frac{12+2}{4} = 3.5$ إما قيمته

فهي متوسط القيمة الثالثة والرابعة أي $Q_1 = \frac{5+7}{2} = 6$. أيضا موقع الربيع الثاني والثالث

هما على التوالي $6.5 = \frac{12+1}{2}$ ، $9.5 = \frac{36+2}{4}$ وعلى ذلك فإن قيمة الربيع الثاني والثالث

على التوالي هما $Q_2 = \frac{11+13}{2} = 12$ ، $Q_3 = \frac{17+19}{2} = 18$.

شكل (١٢-٥)

1 4 5	7 10 11	13 16 17	19 20 22
-------	---------	----------	----------

عند استخدام التوزيعات التكرار يه، فإن قيمة الربيع الأول والثالث تحسب من

المعادلتين:-

$$Q_3 = L + \left(\frac{\frac{3n}{4} - F}{f_{Q_3}} \right) \Delta \quad , \quad Q_1 = L + \left(\frac{\frac{n}{4} - F}{f_{Q_1}} \right) \Delta .$$

حيث L = الحد الأدنى الفعلي لفة الربيع و F = التكرار المتجمع السابق لفة الربيع و Δ = طول لفة الربيع و f_Q = تكرار لفة الربيع .

مثال (١٥-٥) أوجد الربيع الأول والثالث للبيانات في جدول (١٣-٥) .

الحل . موقع الربيع الأول هو $\frac{n}{4} = \frac{100}{4} = 25$ و لفة الربيع الأول هي $59.5 - 49.5$ وقيمة

الربيع الأول هو :-

$$Q_1 = L + \left(\frac{\frac{n}{4} - F}{f_{Q_1}} \right) \Delta = 49.5 + \frac{(25-23)}{16} \cdot 10 = 49.5 + 1.25 = 50.75 .$$

بنفس الشكل موقع الربيع الثالث $\frac{3n}{4} = \frac{(3)(100)}{4} = 75$ و لفة الربيع الثالث هي - 69.5

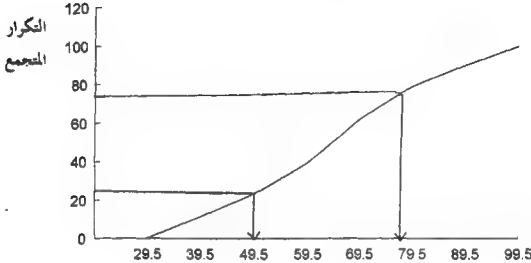
79.5 وقيمة الربيع الثالث هو :-

$$Q_3 = L + \left(\frac{\frac{3n}{4} - F}{f_{Q_3}} \right) \Delta = 69.5 + \frac{(75-62)}{17} \cdot 10 = 69.5 + 7.6471 = 77.1471$$

ويمكن الحصول على الربيع الأول والثالث ببيانات من المضلع التكراري المتجمع بنفس

الطريقة التي استخدمت في حساب الوسيط بيانا مع استخدام موقع الربيع بدلا من موقع الوسيط . باستخدام البيانات المعطاة في جدول (١٣-٥) . فإن قيم الربيع الأول والثالث بيانا هي على التوالي , 49.5 و 77.0 وذلك من شكل (١٣-٥) .

شكل (١٣-٥)



$$Q_1 = 49.5 \quad Q_3 = 77 \quad \text{الحدود العليا للفترة}$$

أيضا يمكن إيجاد القيم التي تقسم فئة من المشاهدات (بعد ترتيبها تصاعديا) إلى عشرة أقسام ونرمز لنقط التقسيم بالرموز D_1, D_2, \dots, D_9 حيث D_1 العشر الأول وهو يمثل القيمة التي يسبقها $\frac{1}{10}$ من البيانات ويلها $\frac{9}{10}$ من البيانات و D_2 العشر الثاني وهو يمثل القيمة التي يسبقها $\frac{2}{10}$ من البيانات ويلها $\frac{8}{10}$ من البيانات وهكذا للعشرات الأخرى . بنفس الشكل يمكن إيجاد القيم التي تقسم فئة من المشاهدات (بعد ترتيبها تصاعديا) إلى مائة قسم ونرمز لنقط التقسيم بالرموز P_1, P_2, \dots, P_{99} حيث P_1 المئين الأول هو يمثل القيمة التي يسبقها $\frac{1}{100}$ من البيانات ويلها $\frac{99}{100}$ من البيانات و P_2 المئين الثاني وهو يمثل القيمة التي يسبقها $\frac{2}{100}$ من البيانات ويلها $\frac{98}{100}$ من البيانات وهكذا لباقي المئينات . في حالة التوزيعات التكرارية يمكن حساب العشرات و المئينات بنفس طريقة حساب الوسيط مع استبدال $\frac{n}{2}$ — $\frac{n}{10}$ للعشر

الأول و $\frac{2n}{10}$ للعدد الثاني وهكذا لباقي العشرات . أيضا استبدال $\frac{n}{2}$ بـ $\frac{n}{100}$ للمئين الأول و $\frac{2n}{100}$ للمئين الثاني وهكذا لباقي المئينات .

مثال (١٦-٥) أوجد العدد الأول والمئين التسعين للبيانات في جدول (١٣-٥) .

الحل . موقع العدد الأول هو $10 = \frac{100}{10} = \frac{n}{10}$ و فئة العدد الأول هي 29.5-39.5 و قيمة العدد الأول هو :-

$$D_1 = 29.5 + \frac{(10-0)}{11} \cdot 10 = 29.5 + 9.091 = 38.591.$$

حيث 29.5 = الحد الأدنى الفعلي لفئة العدد الأول و 0 = التكرار المتجمع السابق لفئة العدد و 10 = طول فئة العدد الأول و 11 = تكرار فئة العدد الأول .

بنفس الشكل موقع المئين التسعين $90 = \frac{(90)(100)}{100} = \frac{90n}{100}$ و فئة المئين التسعين هي 79.5-89.5 و قيمة المئين التسعين هو :-

$$P_{90} = 79.5 + \frac{(90-79)}{11} \cdot 10 = 79.5 + 10 = 89.5.$$

حيث 79.5 = الحد الأدنى الفعلي لفئة المئين التسعين و 79 = التكرار المتجمع السابق لفئة المئين التسعين و 10 = طول فئة المئين التسعين و 11 = تكرار فئة المئين التسعين .

يمكن الحصول على العشرات والمئينات بآيات من المصطلح التكراري المتجمع بنفس الطريقة التي استخدمت في حساب الوسيط بآيات . و مما يجدر الإشارة إليه أن

$Q_2 = D_5 = P_{50}$ ، أيضا كل الربعات والعشرات تمثل مئينات فعلى سبيل المثال D_8 هو P_{80} و Q_3 هو P_{75} .

(١٦-٥) مقاييس التشتت Measures of Dispersion

مقاييس الرعة المركزة التي تمت مناقشتها في البند السابق لا تكفي لإعطاء وصف كافي لتوزيع فئة من المشاهدات فلا توضح طبيعتها ولا كيفية توزيع مشاهداتها . كما أن الاعتماد فقط على أي مقياس للرعة المركزة لمقارنة عدة مجموعات لا يكفي لإظهار حقيقة المقارنة ، فمن الممكن أن يكون لعدة مجموعات من البيانات نفس الوسيط الحسابي والوسيط ولكنهم يختلفون عن بعضهم تمام الاختلاف . فقد تكون مشاهدات إحدى المجموعات متقاربة بعضها من بعض (مركزة حول متوسطها) أو مبعثرة (متشعبة) . فعلى سبيل المثال في جدول (١٥-٥) ثلاثة فئات من المشاهدات A_1, A_2, A_3 كل فئة لها نفس الوسيط الحسابي (60) والوسيط (60)

، ولكن يختلفوا في التشتت أو الانتشار. في الفئة A_1 القيم الستة لهم نفس القيمة ولا يوجد أي تشتت لهم متجانسين تماما. في الفئة A_2 القيم تختلف من خلال ثلاث قيم وفي الفئة A_3 القيم تختلف من خلال قيمتين ولكن هناك انتشار أكثر في الفئة A_3 حيث لا يوجد قيمة في الفئة A_3 تقترب من المتوسط الحسابي أو الوسط. ومن هنا كان من الضروري عند وصف فئة من البيانات بمقياس رقمي أن نصلها عن طريق مقياس من مقاييس الرعة المركزية ومقياس آخر يقيس بعد البيانات عن بعضها أو بعدها عن المتوسط. بعبارة أخرى نصف درجة تشتتها. في الجزء التالي سوف نقدم بعض مقاييس التشتت الأكثر أهمية.

جدول (٥-١٥)

A_1	A_2	A_3
60	35	0
60	35	0
60	60	0
60	60	120
60	85	120
60	85	120

(٥-٦-١) المدى ونصف المدى الربيعي

The range and semi interquartile range

المدى لفئة من المشاهدات، هو الفرق بين أكبر وأصغر مشاهدة. ففي جدول (٥-١٥) المدى للفئة A_1 هو صفر، والمدى للفئة A_2 هو 50 والمدى للفئة A_3 هو 120. وعلى ذلك يعتبر المدى مقياس للتشتت من السهل جانا حسابه ويعطى فكرة سريعة جدا عن طبيعة البيانات ويستخدم كثيرا في مراقبة الجودة وكذلك في وصف الأحوال الجوية. ولكن من عيوبه أنه يتأثر بالقيم الشاذة ويعطى معلومات خاطئة عن الانتشار الحقيقي لمعظم البيانات كما أنه لا يستخدم جميع البيانات في حسابه.

عند استخدام التوزيعات التكرارية فإن المدى يحسب بعدة طرق سوف نذكر منها

الطريقة التالية :

المدى = الحد الأعلى الفعلي للفئة الأخيرة - الحد الأدنى الفعلي للفئة الأولى.

للبيانات في جدول (٥-١٣) المدى = $99.5 - 29.5 = 70$.

هناك مقاييس أخرى للتشتت يمكن استخدامها بدلا من المدى في حالة وجود قيم شاذة.

تعتمد هذه المقاييس على إهمال جزء من البيانات عند طرقي التوزيع حتى نتخلص من القيم الشاذة وتسمى شبيهات المدى. فمثلا يحذف أعلى 10% من المشاهدات وأصغر 10% منها

نحصل على المدى المثيق أي $P_{90} - P_{10}$ • هذا المقياس يستخدم في اختبارات الذكاء في مجال التربية وعلم النفس • أيضا يحدف أعلى 25% من البيانات وأصغر 25% منها نحصل على المدى الربيعي أي $Q_3 - Q_1$ • وأخيرا هناك مقياس آخر يستنتج من المدى الربيعي وهو نصف المدى الربيعي (الانحراف الربيعي) $semi\ interquartile\ range$ ونحصل عليه بقسمة المدى الربيعي على 2 فإذا رمزنا له بالرمز MR فإن :-

$$MR = \frac{Q_3 - Q_1}{2}$$

وعلى ذلك فإن نصف المدى الربيعي لثقة المشاهدات في شكل (١٢-٥) هو :-

$$MR = \frac{Q_3 - Q_1}{2} = \frac{18 - 6}{2} = 6.$$

نضمت

أيضا المدى الربيعي للتوزيع التكراري في جدول (١٣-٥) هو :-

$$MR = \frac{Q_3 - Q_1}{2} = \frac{77.1471 - 50.75}{2} = 13.198550.$$

يعاب على نصف المدى الربيعي أنه لا يستعمل جميع مشاهدات العينة في حسابه ولتلافي هذا العيب سوف نقدم في البند التالي بعض المقياس الأخرى للعثت التي تستخدم جميع البيانات في حسابها •

١ (٢-٦-٥) الانحراف المتوسط The average Deviation

تمثل $|x_i - \mu|$ أو $|x_i - \bar{x}|$ القيمة المطلقة لانحراف أى قيمة عن الوسط الحسابي للمجتمع أو العينة على التوالي •

تعريف : إذا كانت لدينا الثقة من المشاهدات x_1, x_2, \dots, x_n فإن الانحراف المتوسط يمكن حسابه من الصيغة التالية :-

$$M.D = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

مثال (١٧-٥) أوجد الانحراف المتوسط لثقة المشاهدات : 2,3,5,7,8

الحل . من القيم نجد أن $\bar{x} = 5$ والانحرافات : 3,-2,0,2,3 والقيم المطلقة: 3,2,0,2,3 وعلى ذلك:

$$M.D = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} = \frac{3+2+0+2+3}{5} = 2.$$

تعريف : إذا كانت x_1, x_2, \dots, x_k تمثل مراكز الفئة لتوزيع تكراري مع تكراراتها المقابلة f_1, f_2, \dots, f_k فإن الانحراف المتوسط هو :-

$$M.D = \frac{\sum_{i=1}^k f_i |x_i - \bar{x}|}{\sum_{i=1}^k f_i}$$

مثال (١٨-٥) أوجد الانحراف المتوسط للبيانات في جدول (١٦-٥) .

جدول (١٦-٥)

مركز الفئة x_j	التكرار f_j	$x_j - \bar{x}$	$f_j x_j - \bar{x} $
34.5	11	-29.6	325.6
44.5	12	-19.6	235.2
54.5	16	-9.6	153.6
64.5	23	0.4	9.2
74.5	17	10.4	176.8
84.5	11	20.4	224.4
94.5	10	30.4	304
المجموع	100		1428.8

الحل . الوسط الحسابي سبق حسابه من جدول (١٢-٥) وهو $\bar{x} = 64.1$ وعلى ذلك فإن الانحراف المتوسط من جدول (١٦-٥) هو :-

$$M.D = \frac{\sum_{i=1}^k f_i |x_i - \bar{x}|}{\sum_{i=1}^k f_i} = \frac{1428.8}{100} = 14.288$$

(٣-١٦-٥) التباين The Variance

واحد من خصائص الوسط الحسابي ، كما ذكرنا سابقا ، هو أن مجموع مربعات

انحرافات قيم المشاهدات عن الوسط الحسابي للبينة ، أي $\sum_{i=1}^n (x_i - \bar{x})^2$ ، أقل ما يمكن . وإذا

كان اهتمامنا بالمجتمع فإن المقدار $\sum_{i=1}^N (x_i - \mu)^2$ أيضا يكون أقل ما يمكن ، وعلى ذلك إذا أخذنا

هذا المقدار الأخير وقسمناه على حجم المجتمع فإننا نحصل على تباين المجتمع .

تعريف : إذا أعطيت مجتمع محدود x_1, x_2, \dots, x_N فإن تباين المجتمع هو :-

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

الانحراف المعياري، ويرمز له بالرمز σ يمكن الحصول عليه بأخذ الجذر التربيعي للتباين.

مثال (١٩-٥) أوجد الانحراف المعياري لملاحظات المجتمع ٣, ٤, ٥, ٥, ٦, ٧.

الحل . لتسهيل العمليات الحسابية يمكن استخدام جدول (١٧-٥) للحصول على الوسط

الحسابي كما يأتي:-

$$\mu = \frac{\sum_{i=1}^N x_i}{N} = \frac{30}{6} = 5.$$

أيضا التباين:-

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{6} = \frac{10}{6} = \frac{5}{3}.$$

والانحراف المعياري هو : $\sigma = \sqrt{\frac{5}{3}} = 1.29$

جدول (١٧-٥)

x_i	$(x - \mu)$	$(x - \mu)^2$
3	-2	4
4	-1	1
5	0	0
5	0	0
6	1	1
7	2	4
30	0	10

تباين العينة، يرمز له بالرمز s^2 ، ويمثل قيمة من قيم الإحصاء s^2 . عينات عشوائية مختلفة من

الحجم n مختارة من نفس المجتمع، عموماً، تؤدي إلى قيم مختلفة من s^2 . في كثير من المشاكل

قيمة σ^2 تكون غير معروفة وتقدر بالقيمة s^2 . لكي يكون تقديرنا جيد لابد من حساب التباين من صيغة بحيث في المتوسط ينتج القيمة الحقيقية σ^2 . بمعنى أننا إذا أخذنا كل العينات الممكنة من الحجم n من المجتمع وحصلنا على قيمة s^2 لكل عينة، فإن متوسط كل قيم s^2 لابد أن تساوي σ^2 . الإحصاء الذي يحقق هذا الشرط يسمى إحصاء غير متحيز **unbiased**.

تعريف : إذا سحبت العينة العشوائية x_1, x_2, \dots, x_n فإن تباين العينة الغير متحيز هو :-

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

والانحراف المعياري للعينة هو $s = \sqrt{s^2}$

مثال (٥-٢) أوجد تباين العينة والانحراف المعياري للعينة التي مشاهدتها 12, 15, 17, 20.

الحل : التباين والانحراف المعياري للعينة يمكن حسابه من جدول (٥-١٨)

جدول (٥-١٨)

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
12	-4	16
15	-1	1
17	1	1
20	4	16
64	0	34

حيث :-

الوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{64}{4} = 16.$$

وتباين العينة هو :-

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{34}{3} = 11.33.$$

والانحراف المعياري للعينة هو $\sqrt{11.33} = 3.37$

هناك صيغة أخرى لحساب تباين العينة تفيد عند استخدام الآلة الحاسبة وهي :-

$$s^2 = \frac{1}{n-1} \left[\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right].$$

مثال (٥-٢١) أوجد التباين والانحراف المعياري للعينة التي مشاهدتها 3, 4, 5, 6, 7

الحل . بما أن : $\sum_{i=1}^n x_i = 31$, $\sum_{i=1}^n x_i^2 = 171$ وعلى ذلك :-

$$s^2 = \frac{1}{5} \left[171 - \frac{(31)^2}{6} \right] = \frac{13}{6} = 2.1667.$$

الانحراف المعياري للعينة نحصل عليه بإيجاد الجذر التربيعي لتباين العينة . أي أن

$$s = \sqrt{\frac{13}{6}} = 1.47196.$$

يمتاز الانحراف المعياري عن التباين بأنه يعبر عنه باستخدام نفس وحدات القياس كما

في البيانات بينما يكون تميز التباين " وحدات القياس مربعة " .

عرفنا من نظرية (٤-٤) أن $\sigma_{X+b}^2 = \sigma^2$ وهذا يعني أن التباين لنفسة من

المشاهدات لا يتأثر إذا أضفنا ثابت أو طرحنا ثابت من كل مشاهدة . أيضا عرفنا من نظرية (

٤-٥) أن $\sigma_{aX}^2 = a^2 \sigma^2$ وهذا يعني أنه إذا ضربنا كل مشاهدة في ثابت (أو قسمنا على

ثابت) فإن التباين الأصلي نحصل عليه من التباين الجديد بقسمته على (أو ضربه في) مربع

الثابت .

في حالة التوزيعات التكرارية فإن تباين العينة يحسب من المعادلة التالية :-

$$\left(\sum_{i=1}^k f_i = n \text{ حيث} \right) s^2 = \frac{1}{n-1} \left[\sum_{i=1}^k x_i^2 f_i - \frac{\left(\sum_{i=1}^k x_i f_i \right)^2}{n} \right].$$

مثال (٥-٢) أوجد التباين للملاحظات في جدول (٥-١٩) .

الحل . بالتعويض في صيغة التباين السابقة فإن :-

$$s^2 = \frac{1}{99} \left[442265 - \frac{(6410)^2}{100} \right] = 317.0101.$$

جدول (٥-١٩)

حدود الفئة	مركز الفئة X_i	التكرار f_i	$X_i f_i$	$X_i^2 f_i$
30-39	34.5	11	379.5	13092.75
40-49	44.5	12	534.0	23763.00
50-59	54.5	16	872.0	47524.00
60-69	64.5	23	1483.5	95685.75
70-79	74.5	17	1266.5	94354.25
80-89	84.5	11	929.5	78542.75
90-99	94.5	10	945.0	89302.50
المجموع		100	6410	442265

نظرية شيبشيفي Chebyshev's theorem

هذه النظرية مهمة في وصف فئة من المشاهدات . إذا كانت x_1, x_2, \dots, x_n فئة من المشاهدات فإنه على الأقل $(1 - \frac{1}{k^2})$ من المشاهدات سوف تقع ضمن k انحرافات معيارية من وسطها الحسابي . تعتبر النظرية غير مفيدة عندما k تساوى واحد لأن هذا يعنى أنه على الأقل $(1 - \frac{1}{1^2}) = 0$ من المشاهدات تقع في الفترة من $\mu - \sigma$ إلى $\mu + \sigma$. عندما $k = 2$ فهذا يعنى أنه على الأقل $(1 - \frac{1}{2^2}) = \frac{3}{4}$ من المشاهدات تقع في الفترة من $\mu - 2\sigma$ إلى $\mu + 2\sigma$.

وأخيرا عندما $k = 3$ فهذا يعنى أنه على الأقل $(1 - \frac{1}{3^2}) = \frac{8}{9}$ من المشاهدات تقع في الفترة من $\mu - 3\sigma$ إلى $\mu + 3\sigma$. يفرض أن فئة من المشاهدات تمثل عينة، وإذا كانت $k=2$ فالنظرية تنص على أنه على الأقل $(1 - \frac{1}{2^2}) = \frac{3}{4}$ من المشاهدات تقع في الفترة من $\bar{x} - 2s$ إلى $\bar{x} + 2s$. بنفس الشكل، النظرية تنص على أنه $(1 - \frac{1}{2^3}) = \frac{8}{9}$ من المشاهدات تقع في الفترة من $\bar{x} - 3s$ إلى $\bar{x} + 3s$.

مثال (٥-٣) إذا كان الوسط الحسابي والانحراف المعياري لعينة من المشاهدات هما على التوالي 60 و 10 . أستخدم نظرية تشيبشيفي لوصف توزيع المشاهدات .

الحل . (أ) على الأقل $\frac{3}{4}$ من المشاهدات تقع في الفترة 60 ± 20 ، أى من 40 إلى 80 .

(ب) على الأقل $\frac{8}{9}$ من المشاهدات تقع في الفترة 60 ± 30 ، أى من 30 إلى 90 .

مثال (٥-٢٤) المشاهدات التالية تمثل مشاهدات عينة :

14, 18, 19, 19, 19, 20, 20, 20, 20, 21, 25

ضمن اثنين انحراف معياري من الوسط الحسابي وأنه على الأقل $\frac{8}{9}$ من المشاهدات تقع ضمن

ثلاثة انحراف معياري من الوسط الحسابي .

الحل . الوسط الحسابي والانحراف المعياري هما : $s = 2.468$, $\bar{x} = 19.5$ وعلى ذلك $\bar{x} \pm 2s = 19.5 \pm (2)(2.468) = 19.5 \pm 4.94$. يوضح العد الحقيقي أن الفترة من

14.56 إلى 24.44 تحتوي على 10 مشاهدة من 12 مشاهدة، تمثل تقريبا 83% من المشاهدات، أي أكثر من $\frac{3}{4}$ القمم، أيضا

الفترة من 12.1 إلى 26.9 تحتوي على كل القيم في العينة وبالتأكيد تحتوي على الأقل على $\frac{8}{9}$ منهم.

الآن سوف نقدم قاعدة نصف بدلة الانتشار للتوزيع الناقوس. يعرف التوزيع الناقوس عادة بالتوزيع الطبيعي والذي سوف نناقشه بالتفصيل في الفصل السادس. كما تعطى هذه القاعدة نتائج جيدة للتوزيعات القريبة الشكل من التوزيع الطبيعي.

قاعدة تجريبية Empirical Rule

أعتبر توزيع فئة من المشاهدات لها التوزيع الناقوس، الفترة :-

(أ) $\bar{x} \pm s$ سوف تحتوي تقريبا على 68% من المشاهدات .

(ب) $\bar{x} \pm 2s$ سوف تحتوي تقريبا على 95% من المشاهدات .

(جـ) $\bar{x} \pm 3s$ سوف تحتوي تقريبا على 99.7% من المشاهدات .

مثال (٥-٢٥) المشاهدات في جدول (٥-٢٠) تمثل التوزيع التكراري لدرجات مجموعة من

الطلبة في مادة الإحصاء ، المطلوب إيجاد التوزيع التكراري ووصف البيانات منه .

الحل . للبيانات في جدول (٥-٢١) الوسط الحسابي والانحراف المعياري على التوالي هما

$\bar{x} = 66.1$ ، $s = 13.7$ وعلى ذلك يمكننا حساب الفترات :

(أ) $\bar{x} \pm s = 66.1 \pm 13.7$

(ب) $\bar{x} \pm 2s = 66.1 \pm 27.4$

(جـ) $\bar{x} \pm 3s = 66.1 \pm 41.1$

جدول (٥-٢٠)

35	41	44	45	40	51	48
54	56	55	53	58	59	60
60	61	62	63	67	64	64
67	65	66	68	69	66	70
73	75	74	72	71	76	81
79	80	78	82	83	85	86
50	62	68	72	80	88	51
91	92					

ال تكرار	حدود الفئة
1	35-39
3	40-44
2	45-49
5	50-54
4	55-59
8	60-64
8	65-69
6	70-74
4	75-79
5	80-84
3	85-89
2	90-94

جدول (٥-٢١)

القاعدة التجريبية السابقة سوف نعيدنا كثيرا في وصف المشاهدات في جدول (٥-٢١) . تبعا لهذه القاعدة فإننا نتوقع أن تقريبا 68% من المشاهدات تقع في الفترة 52.4 إلى 79.8 . يوضح العد الحقيقي من جدول (٥-٢٠) أن 34 من 51 مشاهدة تقع في هذه الفترة وتمثل 67.0% من المشاهدات . أيضا نتوقع أن تقريبا 95% من المشاهدات تقع في الفترة 38.7 إلى 93.5 . يوضح العد الحقيقي أن 49 من 51 مشاهدة تقع في هذه الفترة وتمثل 96% من المشاهدات . وأخيرا نتوقع أن تقريبا 99.7% من كل المشاهدات تقع في الفترة 25 إلى 107.2 ، العد الحقيقي يوضح أن كل المشاهدات تقع في هذه الفترة وتمثل 100% من كل المشاهدات .

Coefficient of Variation

(٥-٦-٤) معامل الاختلاف

تعتبر كل مقاييس الششت السابقة مقاييس مطلقة لأنها تأخذ تميز الوحدات الأصلية ولذلك لا تصلح للمقارنة بين مجموعتين وحدات القياس بينهما مختلفة . لذلك سوف نناقش مقياس نسبي يسمى معامل الاختلاف والذي يحول الانحراف المعياري إلى مقياس نسبي باعتبار أنه نسبة متوية من الوسط الحسابي . ويمكن حساب معامل الاختلاف من إحدى المعادلتين التاليتين :

$$V = \left(\frac{\sigma}{\mu} \right) 100 \quad \text{أو} \quad V = \left(\frac{s}{\bar{x}} \right) 100 .$$

لتسهيل استخدام معامل الاختلاف نقدم المثال التالي والذي يوضح كمية الإنتاج في شركة ما خلال 80 شهرا ثم كمية الإنتاج خلال فترة ثالثة مقدارها 15 شهرا وقد تم حساب الوسط الحسابي والانحراف المعياري ومعامل الاختلاف لكل مجموعة والنتائج في جدول (٥-٢٢) . يلاحظ من النتائج أن الإنتاج خلال 15 شهرا له وسط حسابي أكبر ومعامل اختلاف أقل والذي يعتبر نتائج جيدة لمدير الإنتاج والذي يهتم بزيادة الإنتاج وانخفاض معامل الاختلاف . وبالرغم

من أن الانحراف المعياري قد زاد من 13.2 إلى 15 إلا أنه يمكن القول بناء على معامل الاختلاف أن الفترة الثانية أقل تشعباً من الفترة الأولى .

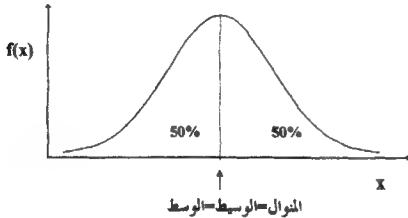
جدول (٥-٢٢)

الفترة	\bar{x}	s	$V = \left(\frac{s}{\bar{x}}\right) 100$
80	125	13.2	10.56
15	160	15	9.375

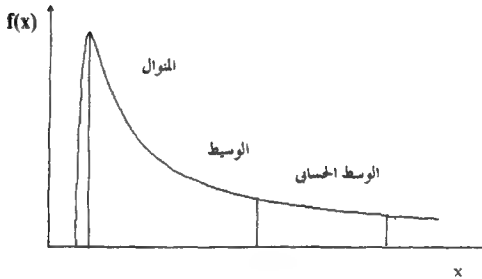
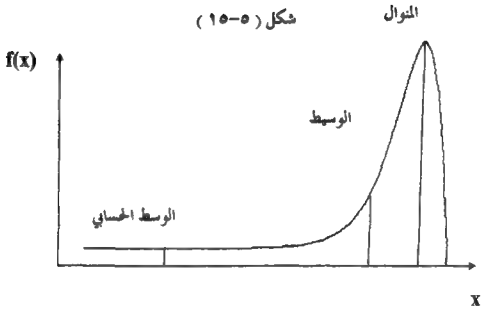
(٥-٧) الالتواء والعلاقة بين الوسط الحسابي والوسيط والمنوال

Skewness and the Relation of the Mean , Median , and Mode

عرفنا مما سبق أن الالتواء هو بعد التوزيع التكراري عن التماثل . فإذا كان التوزيع متماثلاً فسوف نجد أن 50% من القيم تقع على كل جانب من المنوال كما في شكل (٥-١٤) .



أيضاً نلاحظ من شكل (٥-١٤) أن التوزيع له منوال واحد unimodal (وحيد المنوال) وأن الوسط الحسابي = الوسيط = المنوال . بينما في شكل (٥-١٥) نجد أن هناك علاقة بين الوسط الحسابي والوسيط والمنوال حيث الوسط الحسابي > الوسيط > المنوال وذلك لأن التوزيع ملتويًا جهة اليسار . بينما في شكل (٥-١٦) نجد أن الوسط الحسابي < الوسيط < المنوال وذلك لأن التوزيع ملتويًا جهة اليمين . وفي كلتا الحالتين فإن الوسيط يقع بين الوسط الحسابي والمنوال كما أن الوسط الحسابي يقع دائماً في اتجاه القيم الشاذة .



شکل (١٦-٥)

(٨-٥) بعض مقاييس الالتواء وانفطاح

Some Measures of Skewness and Kurtosis

أولاً بالنسبة لمقاييس الالتواء ، سوف نتناول مقياسين للالتواء الأول ويسمى معامل بيرسون للالتواء Pearsonian coefficient for skewness . تعرف معادلة معامل بيرسون للالتواء كالتالي :-

$$S_k = \frac{3(\bar{x} - \tilde{x})}{s}.$$

حيث \bar{x} الوسط الحسابي و s الانحراف المعياري للعينة. ينحصر قيمة معامل بيرسون بين -3 إلى +3 ، عندما $S_k = 0$ فهذا يعني أن التوزيع متماثل ، وإذا كانت قيمة S_k موجبة فهذا يعني أن الوسط الحسابي أكبر من الوسيط ومن النوال وبذلك يكون المنحنى ملتويا وله ذيل ناحية اليمين ويكون الالتواء موجبا . وأخيرا وإذا كانت قيمة S_k سالبة فهذا يعني أن الوسط الحسابي أصغر من الوسيط ومن النوال وبذلك يكون المنحنى ملتويا وله ذيل ناحية اليسار ويكون الالتواء سالبا . من جدول (٥-٢٢) وإذا كان الوسيط للبيانات في الفترة التي تمثل 80 شهرا هي 115.3 فإن معامل الالتواء هو :-

$$S_k = \frac{3(\bar{x} - \tilde{x})}{s} = \frac{3(125 - 115.3)}{13.2} = +2.2045.$$

وهذا يعني أن التوزيع به كمية من الالتواء الموجب .

يعتمد المقياس السابق للالتواء على أنه في التوزيعات المتلوية فإن الوسيط يقع تقريبا

في $\frac{1}{3}$ المسافة بين الوسط الحسابي والنوال في اتجاه الوسط الحسابي كما في شكل (٥-١٥) وشكل (٥-١٦) وهذا غير صحيح دائما . ولذلك سوف نتناول مقياس آخر للالتواء يعتمد على العزم المقدر من بيانات العينة .

تعريف : العزم r حول المتوسط لفترة المشاهدات x_1, x_2, \dots, x_n هو :-

$$m^r = \frac{\sum_{i=1}^n (x_i - \bar{x})^r}{n}.$$

في حالة التوزيعات التكرارية وإذا كانت x_1, x_2, \dots, x_k تمثل مراكز الفئات

لتوزيع تكراري مع تكراراتها المقابلة f_1, f_2, \dots, f_k (حيث k تمثل عدد الفئات) فإن العزم

r يحسب من الصيغة التالية :-

$$m^r = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^r}{\sum_{i=1}^k f_i}.$$

المقياس الثاني للالتواء والذي يعتمد على العزم الثالث حول الوسط الحسابي هو:

$$a_1 = \frac{m^3}{s^3}.$$

إذا كان التوزيع متماثل ، فهذا يدل على أن $a_1 = 0$ وإذا كان $a_1 > 0$ يكون التوزيع موجب الالتواء . وإذا كان $a_1 < 0$ يكون التوزيع سالب الالتواء .
 ثانياً بالنسبة لمقاييس التفلطح سوف نتناول مقياس يعتمد على العزم الرابع حول المتوسط معادلته هي :-

$$a_2 = \frac{m^4}{s^4}.$$

إذا كانت $a_2 = 3$ ، فهذا يعني أن التوزيع متوسط التفلطح . وإذا كان $a_2 > 3$ ، فهذا يعني أن التوزيع له قمة مدببة وإذا كان $a_2 < 3$ فهذا يدل على أن التوزيع مفلطحاً .
 مثال (٢٦-٥) أوجد مقياس الالتواء a_1 ومقياس التفلطح a_2 لفئة المشاهدات 2,4,6,8,13,15 .

الحل . الجدول (٢٣-٥) يسهل عملية الحساب كالتالي :-

جدول (٢٣-٥)

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
2	-6	36	-216	1296
4	-4	16	-64	256
6	-2	4	-8	16
8	0	0	0	0
13	5	25	125	625
15	7	49	343	2401
48	0	130	180	4594

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{130}{5} = 26 , \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{48}{6} = 8$$

$$s^3 = (26)(5.09902) = 132.5745 , \quad s^4 = 676 , \quad \sum_{i=1}^n (x_i - \bar{x})^3 = 180$$

وعلى ذلك العزم الثالث حول الوسط الحسابي هو :-

$$m^3 = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n} = \frac{180}{6} = 30.$$

ومقياس الالتواء :-

$$a_1 = \frac{m^3}{s^3} = \frac{30}{132.5745} = 0.226288.$$

ويمكن إيجاد مقياس التفلطح a_2 وذلك بحساب القيم التالية من جدول (٢٣-٥) :-

$$m^4 = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n} = \frac{4594}{6} = 765.667, \quad \sum_{i=1}^n (x_i - \bar{x})^4 = 4594.$$

وعلى ذلك نحصل على مقياس التفلطح :-

$$a_2 = \frac{m^4}{s^4} = \frac{765.667}{676} = 1.13264.$$

وهذا يدل على أن توزيع المشاهدات مفطح.

مثال (٢٧-٥) أوجد مقياس للتفلطح ومقياس للالتواء للمشاهدات في جدول (٢٤-٥) .

جدول (٢٤-٥)

الحدود الفئة	2-6	7-11	12-16	17-21	22-26	المجموع
مركز الفئة x_i	4	9	14	19	24	
التكرار f_i	2	3	5	3	7	20
$x_i f_i$	8	27	70	57	168	330
$(x_i - \bar{x})$	-12.5	-7.5	-2.5	2.5	7.5	
$(x_i - \bar{x}) f_i$	-25	-22.5	-12.5	7.5	52.5	
$(x_i - \bar{x})^2 f_i$	312.5	168.75	31.25	18.75	393.75	925
$(x_i - \bar{x})^3 f_i$	-3906.3	-1265.6	-78.1	46.9	2953.1	-2250
$(x_i - \bar{x})^4 f_i$	48828.8	9492.0	195.3	117.3	22148.3	80781.7

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i} = \frac{330}{20} = 16.5$$

الحل . من جدول (٢٤-٥) فإن الوسط الحسابي هو :

والتباين يساوي :

$$s^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 f_i}{(\sum_{i=1}^k f_i - 1)} = \frac{925}{19} = 48.68421.$$

$$s^3 = 339.6896, s^4 = 2370.1523, \sum_{i=1}^k (x_i - \bar{x})^3 f_i = -2250.$$

$$m^3 = \frac{\sum_{i=1}^k (x_i - \bar{x})^3 f_i}{(\sum_{i=1}^k f_i)} = \frac{-2250}{20} = -112.5.$$

وعلى ذلك يمكن الحصول على مقياس الالتواء كما يلي :-

$$a_1 = \frac{m^3}{s^3} = \frac{-112.5}{339.6896} = -0.331185.$$

ويمكن إيجاد مقياس التفلطح a_2 وذلك من جدول (٥-٢٤):

$$\sum_{i=1}^k (x_i - \bar{x})^4 f_i = 80781.7,$$

$$m^4 = \frac{\sum_{i=1}^k (x_i - \bar{x})^4 f_i}{\sum_{i=1}^k f_i} = \frac{80781.7}{20} = 4039.085.$$

وعلى ذلك يمكن الحصول على مقياس للتفلطح كما يلي :-

$$a_2 = \frac{m^4}{s^4} = \frac{4039.085}{2370.1523} = 1.704146.$$

تمارين

- ١ - إذا أعطيت التوزيع الاحتمالي التالي :

x	1	2	3
P(X=x)	0.25	0.25	0.5

(أ) أوجد قيم فئة المشاهدات من الحجم $N = 100$ التي مثلت هذا التوزيع .

(ب) أوجد الوسط الحسابي للـ المجتمع بطريقتين .

٢ - أوجد المعلمة μ للمجتمع 4,6,5 ثم ضع قائمة بكل العينات الممكنة من الحجم $n = 2$ مع إمكانية تكرار القيمة الواحدة في العينة ، واحسب لكل عينة القيمة \bar{X} للإحصاء \bar{X} وأثبت أن $E(\bar{X}) = \mu$.

٣ - في تمرين ٢- أحسب s^2 لكل عينة وذلك تحت فرض (١) μ معروفة (ب) μ غير معروفة وأثبت أن $E(S^2) = \sigma^2$ لكل حالة .

٤ - المطلوب إلقاء 10 عملات 100 مرة وتسجيل قيم x التي تمثل عدد مرات ظهور الصورة ثم إيجاد المدرج التكراري للتوزيع التكراري الذي يمثل عدد مرات ظهور الصورة .

٥ - أوجد للفئات التالية الحدود الفعلية للفئة ومركز الفئة وطول الفئة :

(١) 8-12 (ب) 11.5-16.7 (جـ) 2.2-3.3 (د) 0.345-0.416 .

(ز) 77.45-86.12 .

٦ - فيما يلي التوزيع التكراري للدرجات التي حصل عليها 80 شخصا في دورة تدريبية في مجال الحاسب الآلي استمرت 6 شهور والمطلوب إيجاد : (١) عدد الفئات (ب) طول الفئة (جـ) مركز الفئة (د) التكرار النسبي والتكرار المئوي .

حدود الفئة	50-59	60-69	70-79	80-89	90-99
التكرار	5	10	50	10	5

٧ - إذا كانت مراكز الفئات للتوزيع التكراري الذي يمثل أوزان مجموعة من النباتات (بالرطل) هي : 17,20,23,26,29,32,35 . أوجد حدود الفئة والحدود الفعلية وطول الفئة لهذا التوزيع .

٨- في عينة من الزجاجات سعة كل منها لترا واحدا تم قياس ما تحتويه من سائل بالمليتر وتم وضع البيانات في الجدول التالي:-

حدود الفئة	900-909	910-919	920-929	930-939	940-949
التكرار	5	8	8	24	15

أوجد : الوسط الحسابي لما تحتويه الزجاجات من سائل .

٩ - في عينة عشوائية من 20 طالب في كلية ما تم تسجيل عدد أيام الغياب لكل طالب خلال الفصل الدراسي الأول وكانت كالتالي : 1,0,3,4,5,4,1,1,1,1,0,0,0,3,3,2,2,1 .

أحسب كلا من :

الوسط الحسابي - الوسط - النوال - الانحراف المتوسط - الانحراف المعياري .

- ١٠ - فيما يلي السرعة (بالأميال لكل ساعة) التي سجلها وادار على الطريق الزراعي لعينة عشوائية من 50 سيارة مرت عند نقطة المراقبة خلال ساعة :

60	66	54	54	49	74	71	65	56	47
59	70	71	66	65	70	65	64	63	55
55	54	60	63	61	65	45	53	54	61
54	54	53	48	47	70	74	63	62	61
54	66	70	64	65	64	63	68	66	70

- (أ) ضع هذه البيانات في شكل توزيع تكراري .
 (ب) أوجد التوزيعات التكرارية النسبية والنسبة والمتوية والمتجمعة .
 (جـ) ارسم المدرج التكراري والمضلع التكراري لهذا التوزيع .
 (د) ارسم المنحنى التكراري النسبي والمنحنى التكراري المتجمع النسبي .
 - ١١ - إذا كان عدد أسياتك السالون التي تم صيدها بواسطة 10 صيادين في اليوم الأول من الموسم هي : 5,6,7,7,7,8,9,10,3,7 أوجد : الوسط - الانحراف المتوسط - الانحراف المعياري .

- ١٢ - الجدول التالي يمثل التوزيع التكراري لأعمار مجموعة من أبناء طلبة كلية

حدود الفئة	35-39	40-44	45-49	50-54	55-59
التكرار	6	15	40	25	14

والمطلوب :

- (أ) رسم المدرج التكراري والمضلع التكراري لهذا التوزيع .
 (ب) إيجاد التكرار النسبي لهذا التوزيع ورسم مضلعة التكراري .
 (جـ) إيجاد التكرار المتوي لهذا التوزيع ورسم مضلعة التكراري .
 (د) إيجاد الوسط الحسابي والانحراف المعياري .
 (ز) ما نسبة الآباء الذين أعمارهم : ضمن انحراف معياري واحد من الوسط الحسابي أي الواقعة في الفترة $\bar{x} \pm s$ ؟

(ر) كور (ز) للفترة $\bar{x} \pm 2s$ وأيضا للفترة $\bar{x} \pm 3s$.

- ١٣ - اخبرت عينة عشوائية من 10 مسامير لتقدير كمية الضغط الضروري لكسر المسامير وكانت النتائج كالتالي : 18,22,26,25,27,26,19,17,22,20 : أحسب كلا من :
 الوسط الحسابي - الوسط - المنوال - الانحراف المتوسط - الانحراف المعياري - مقياس الالتواء a_1 ومقياس التغلطح a_2 - مقياس الالتواء لبيرسون .

- ١٤ - فيما يلي التوزيع التكراري لعدد الدقائق التي يتأخرها مجموعة من الطلبة في دخولهم المحاضرة بعد دخول الأستاذ :

التأخير بالدقائق	0	1	2	3	4	5	6
التكرار	180	1	2	3	5	6	6

أوجد الوسط الحسابي والانحراف المعياري لهذا التوزيع .

- ١٥ - تمتلك شركة ما 10 قوارب للصيد، قامت الشركة بتسجيل تكاليف صيانة كل قارب (بالدولار) وكانت كما يلي : 500,505,460,470,530,506,994,880,600,460
أوجد :

(أ) الوسط والوسط الحسابي والنوال . (ب) الربيع الأول والربيع الثالث .

(ج) الانحراف الربيعي والمدى .

(د) مقياس للالتواء وآخر للضلع .

- ١٦ - الجدول التالي يمثل التوزيع التكراري لتكاليف تجهيز عينة عشوائية من نبات للتصدير .

حدود الفئة	1.00-1.02	1.03-1.05	1.06-1.08	1.09-1.11	1.12-1.14
التكرار	5	25	57	40	41

المطلوب إيجاد (أ) الوسط والوسط الحسابي والنوال . (ب) الربيع الأول والربيع الثالث .

- ١٧ - قامت شركة متخصصة في بيع البرامج الجاهزة بتسجيل حصيلة البيع الشهري بالدولار وذلك في مدة 28 شهرا وكانت المشاهدات كما يلي :

757.34	990.16	118.01	871.54	858.29	820.54	710.99
1150.5	1280.3	723.06	876.09	1230.9	657.98	1018.6
1140.6	997.05	657.90	999.98	1140.8	800.75	345.89
1234.8	1280.3	723.06	887.09	1209.0	670.01	678.98

(أ) كون توزيع تكراري مناسب (ب) أوجد الوسط الحسابي لهذا التوزيع .

(ج) أوجد الوسط لهذا التوزيع حسابيا وبيانيا .

(د) ارسم المدرج التكراري وأحسب من النوال .

- ١٨ - إذا كانت درجة الحرارة المتوىة في إحدى المدن خلال أيام إحدى الأسابيع هي :

22,9,13,12,18,15,9 أحسب ما يلي : (أ) الوسط والوسط الحسابي والنوال .

- ١٩ - فيما يلي 25 قيمة لتغير عشوائي :-

34	50	44	54	33
34	32	34	27	22
25	12	60	50	13
40	34	35	35	34
23	34	6	7	18

(١) كون توزيع تكراري مناسب.

(ب) ارسم المدرج التكراري والمضلع التكراري لهذا التوزيع.

(ج) أوجد العزوم الأربعة الأولى حول الوسط الحسابي.

- ٢٠ - مجتمع له وسطه الحسابي $\mu = 11$ وانحرافه المعياري $\sigma = 2$ باستخدام نظرية

تشبيش أوجد (١) نسبة القيم التي تقع بين 7 و 15 (ب) نسبة القيم التي تقع بين 5

و 17.

- ٢١ - فيما يلي التوزيع التكراري للضرائب التي تم تحصيلها من مجموعة من الموظفين في

شركة لتسخين البترول في عام 1990.

حدود الفئة	400-500	501-601	602-702	703-803	804-904
التكرار	17	25	29	25	28

(١) ارسم المضلع التكراري المتجمع ثم من الرسم حدد الوسط والربيع الأول والربيع الثالث.

(ب) أحسب مقياس للدرجة المركبة.

(ج) ما هو الرقم الذي يقسم التوزيع بحيث يكون عدد العاملين التي الضرائب المستحقة

عليهم الأكبر منه يساوي عدد العاملين الأقل منه ؟

(د) ما نسبة العاملين الذين يدفعون ضرائب أقل من 702.5.

- ٢٢ - المشاهدات التالية تمثل التوزيع التكراري للأجور الشهرية (بالدولار) لمجموعة من

العمال في شركة لصناعة الغاز الطبيعي.

الأجور	التكرار
900-950	9
951-1001	15
1002-1052	21
1053-1103	25
1104-1154	27
1155-1205	19
1206-1256	14
1257-1307	9
1308-1358	6

(١) كون توزيع تكراري مناسب .

(ب) ارسم المدرج التكراري والمضلع التكراري لهذا التوزيع .

(ج) أوجد العزوم الأربعة الأولى حول الوسط الحسابي .

-٢٣- الآتي يمثل أعمار 30 عاملة في مصنع لصناعة الحلوى :

43	58	21	20	30	48
50	32	61	29	34	20
24	49	50	31	32	25
23	21	47	34	35	41
34	36	34	32	33	42

المطلوب إيجاد التوزيع التكراري و رسم المدرج و المضلع و المنحنى التكراري .

-٢٤- الجدول التالي يمثل التوزيع التكراري لأعمار مجموعة من الذكور والإناث في كلية ما .

العمر	15-16	17-18	19-20	21-22	23-24
الإناث	100	125	130	160	190
الذكور	110	131	150	146	200

(١) ارسم المضلع التكراري لكل من الذكور والإناث .

(ب) أوجد الوسط الحسابي والانحراف المعياري لكل من الذكور والإناث .

(ج) أوجد معامل الاختلاف لكل من الذكور والإناث وأي المجموعتين أكثر تشتتاً .

-٢٥- إذا كان معدل التضخم لشعب ما هو 3% في السنة الأولى و4% في السنة الثانية

و8% في السنة الثالثة . أوجد الوسط الهندسي لمعدلات التضخم .

-٢٦- حصل مستثمر على عائد على رأس ماله قدره 3% للسنة الأولى و5% للسنة الثانية

12% للسنة الثالثة أوجد الوسط الهندسي للعائد السنوي .

-٢٧- يشتري شخص 4 قمصان من الشركة A بسعر الواحد \$22 و4 قمصان من الشركة

B بسعر الواحد \$25 و7 قمصان من الشركة C بسعر الواحد \$30 . أوجد متوسط سعر

القميص .

-٢٨- أسرة لديها 8 أطفال، أعمارهم كالتالي : 8,10,6,14,14,12,18,20 : أوجد :

(١) مقاييس الرعة المركزية لهذه البيانات .

(ب) المدى - المدى الربيعي - الانحراف المتوسط - الانحراف المعياري .

- ٢٩ - اشترى شخص 10 كيلو من السمك من نوع (A) بسعر الكيلو 4 جنيهها و 5 كيلو من السمك من نوع (B) بسعر الكيلو 2.5 جنيهها و 3 كيلو من السمك من نوع (C) بسعر الكيلو 1.5 . باستخدام الوسط الحسابي المرجح أوجد متوسط سعر السمك الذي اشترى به .

- ٣٠ - يقوم رجل أعمال بتأجير الشقق التي يمتلكها شهريا بأسعار مختلفة . فالشقة من النوع الممتاز بسعر 1500 جنيهها والشقة من النوع الفاخر بسعر 1000 جنيهها والشقة من النوع المتوسط بسعر 500 جنيهها . فإذا كان الرجل يمتلك 50 شقة من النوع الممتاز و 30 شقة من النوع الفاخر و 15 شقة من النوع المتوسط . أوجد الوسط الحسابي للإيجار الشهري الذي يتحصل عليه .

- ٣١ - لدى رجل أعمال أربعة حسابات في بنك ما . ففي الحساب A له 4,000 دولار وفي الحساب B له 3,000 دولار وفي الحساب C له 10,000 دولار . فإذا كان الحساب A يعطى له أرباح قدرها 5.5% كل سنة والحساب B يعطى له أرباح قدرها 4% كل سنة والحساب C يعطى له أرباح قدرها 6% كل سنة . أوجد الوسط الحسابي المرجح للربح السنوي .

- ٣٢ - فيما يلي توزيع عدد المشاريع المنفذة شهريا خلال عام 1995 في شركة بتروك :
15,11,7,6,8,10,12,6,8,9,6,13 : أحسب :

(أ) الوسط الحسابي - الوسيط - المنوال - الانحراف المعياري - المدى - المدى الربيعي .

- ب - مقياس الالتواء a_1 ومقياس للتضلع a_2 - مقياس الالتواء لبيرسون .

- ٣٣ - يقوم العاملين في شركة صغيرة بالتوقيع في صفحات الزمن للدلالة على زمن مفادرة الشركة . المشاهدات التالية تمثل أزمنة المفادرة للعاملين وذلك في يوم عشوائي :

5:15	3:50	1:10	5:40	5:30	5:33
4:45	5:59	4:30	5:12	5:16	4:30
5:30	2:40	5:40	3:30	3:33	3:30
5:30	4:56	3:22	5:50	5:41	4:30
5:30	4:23	4:32	5:44	5:35	5:12

(أ) أوجد الوسط الحسابي لهذا التوزيع .

(ب) أوجد الوسيط لهذا التوزيع .

(ج) أوجد الانحراف المعياري لهذا التوزيع .

- ٣٤ - فيما يلي توزيع عدد مرات احتراق الطعام خلال أسبوع لعينة من 10 سيدات :

3,2,4,1,1,0,0,0,4,1 المطلوب إيجاد الوسط الحسابي والانحراف المعياري ووصف البيانات

باستخدام نظرية تشيبيشيف .

- ٣٥ - الجدول التالي يمثل التوزيع الشكراري للودائع في بنك ما في نهاية عام 1993 :

حجم الودائع	عدد العملاء
0-500	7100
501-1001	8300
1002-1502	9345
1503-2003	9945
2004-2504	6257
2505-3005	2003
3006-3506	14
3507-4007	1445

أوجد : الوسط الحسابي - المتوال - الانحراف المتوسط - الانحراف المعياري

- ٣٦ - إذا كانت جامعة ما تمثل مجتمع عناصره عدد المدرسين في كل كلية . بفرض أن الوسط الحسابي لعدد المدرسين في الكلية الواحدة هو $\mu = 16$ بالانحراف المعياري $\sigma = 15$. استخدم نظرية تشيبيشيف في وصف هذا المجتمع .

- ٣٧ - أجرى اختبار في مادة الإحصاء لمجموعة من الطلبة عددهم 400 وقد وجد أن الوسط الحسابي هو $\bar{x} = 77$ والانحراف المعياري $\sigma = 5$ استخدم نظرية تشيبيشيف لوصف هذه الفئة من البيانات .

- ٣٨ - وجد أن تلوث البحار يؤدي إلى نمو أنواع مختلفة من البكتريا . فإذا كان عدد البكتريا لكل 100 مليلتر في 10 أماكن من مياه البحر هي : 49,69,60,41,69,70,51,68,67,66 : أوجد : الوسط الحسابي - الوسط - المتوال - الانحراف المتوسط - الانحراف المعياري - مقياس للالتواء وآخر للتقلطح .

- ٣٩ - تم سؤال عينة عشوائية من 10 عمال عن المسافة (بالأميال) التي يقطعونها في الذهاب إلى المزرعة التي يعملون بها وكانت إجاباتهم كما يلي : 25,6,1,2,4,8,5,6,5,4 : أوجد : الوسط والربيع الأول والربيع الثالث والمدى الربيعي - مقياس للالتواء وآخر للتقلطح .

- ٤٠ - المشاهدات التالية تمثل عدد المرضى الذين يتم الكشف عليهم يوميا في مستشفى خاص من قبل 10 أطباء : 15,8,6,9,15,18,21,39,5,7 : أوجد الوسط الحسابي والانحراف المعياري واستخدم نظرية تشيبيشيف لوصف هذه الفئة من المشاهدات وأوجد مقياس للالتواء وآخر للتقلطح .

- ٤١ - في جامعة ما يتقاضى الأستاذ في المتوسط \$15,000 دولار بانحراف معياري \$5,000 بينما في جامعة أخرى يتقاضى الأستاذ أجر قدره \$10,000 بانحراف معياري \$3,000 أوجد معامل الاختلاف لكل جامعة وأي الجامعات أكثر تشتتاً ؟

- ٤٢ - تمتلك شركة أربع مزارع ، بعض البيانات عن هذه المزارع في الجدول التالي :

الانحراف المعياري	متوسط الأجر السنوي	عدد العمال	المزرعة
\$75	\$8,300	100	1
1200	11,200	180	2
900	9,000	200	3

أوجد معامل الاختلاف لكل مزرعة وما هي المزرعة التي لها أكبر تشتت في الأجر السنوي ؟

- ٤٣ - إذا كان الوسط الحسابي للأجر الشهري للموظف في شركتين A , B متساوي وهو \$10,000 بانحراف معياري \$500 للشركة A وانحراف معياري \$550 للشركة B . أوجد معامل الاختلاف لكل شركة وأي الشركتين أكثر تشتتاً في الأجر ؟

- ٤٤ - قام مزارع بوزن نوعين من ثمار البرتقال في المزرعة التي يمتلكها وقد حصل على المشاهدات التالية : $\bar{x}_1 = 15$, $s_1 = 1$, $\bar{x}_2 = 14$, $s_2 = 2$ أوجد معامل الاختلاف لكل نوع وأي النوعين أكثر تشتتاً ؟

الفصل السادس

بعض التوزيعات الاحتمالية

Some Probability Distributions

بالرغم من وجود أنواع لانهائية من التوزيعات الاحتمالية فإن عدد محدود منهم يستخدم في مجال واسع من التطبيقات الإحصائية . يتناول هذا الفصل بعض هذه التوزيعات .

(١-٦) التوزيع المنتظم Uniform Distribution

يعتبر التوزيع المنتظم من أبسط التوزيعات الاحتمالية المنفصلة، حيث أن جميع قيم المتغير العشوائي X لها الاحتمال نفسه . لهذا التوزيع بعض التطبيقات المحدودة خاصة في المعاينة الإحصائية .

تعريف : إذا كان فراغ المتغير العشوائي X هو $\Omega = \{x_1, x_2, \dots, x_c\}$ فإن التوزيع المنتظم يأخذ الصيغة التالية :

$$f(x; c) = \frac{1}{c}, \quad x = x_1, x_2, \dots, x_c.$$

سوف نستخدم الصيغة $f(x; c)$ بدلا من $f(x)$ لتوضيح أن التوزيع المنتظم يعتمد على المعلمة c . مثال (١-٦) يتكون الكتاب الخاص بدائرة المعلومات البريطانية لعام ٢٠٢٠ من 20 جزء، فإذا كان المطلوب اختيار جزءاً عشوائياً . أوجد التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل رقم الجزء المختار .

الحل : عند اختيار جزءاً عشوائياً من 20 جزء فإن كل عنصر في فراغ المتغير العشوائي $\Omega = \{x_1, x_2, \dots, x_{20}\}$ يقع باحتمال $\frac{1}{20}$ وعلى ذلك يكون لدينا توزيع منتظم بدالة كثافة احتمال :

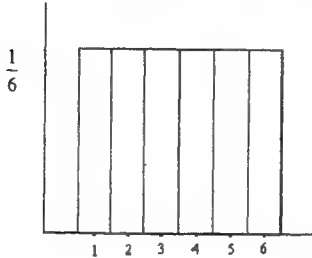
$$f(x; 20) = \frac{1}{20}, \quad x = 1, 2, \dots, 20.$$

مثال (٢-٦) أوجد التوزيع الاحتمالي لكل العينات الممكن اختيارها من الحجم $n = 2$ من القيم $\{1, 2, 3, 4\}$.

الحل : عدد العينات الممكن اختيارها هو $\binom{4}{2} = 6$ وهى على التوالي $\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}$. كل العينات السابقة لها نفس القرصة في الظهور عند اختيار عينة عشوائية من الحجم $n = 2$ من القيم $\{1, 2, 3, 4\}$. توزيع العينات يتبع التوزيع المنتظم بدالة كثافة احتمال :

$$f(x; 6) = \frac{1}{6}, \quad x = 1, 2, \dots, 6.$$

حيث 1 تعني العينة {1,2} و 2 تعني العينة {1,3}... الخ • وعلى ذلك فإن احتمال اختيار العينة {1,2} هو $P(\{1,2\}) = P(X=1) = \frac{1}{6}$ • عموماً عند اختيار عينة عشوائية من الحجم n من مجتمع محدود حجمه N فإن قيمة c في صيغة التوزيع المنتظم تعطى من $\left(\frac{N}{n}\right)$ • عند التمثيل البياني للتوزيع المنتظم بالمرج histogram نحصل على مستطيلات متساوية الارتفاع كما في شكل (١-٦) للمثال (٢-٦) •



شكل (١-٦)

Binomial Distribution توزيع ذي الحدين (٢-٦)

في كثير من الأحيان قد تشتمل تجربة ما على n من المحاولات المتكررة المستقلة بحيث يكون لكل محاولة نتيجتين التين فقط، تسمى الأولى نجاح وتسمى الأخرى فشل، حيث احتمال النجاح p واحتمال الفشل $q = 1 - p$ • تسمى التجربة التي تحقق هذه الشروط بتجربة نسميها الحدين binomial experiment • فعلى سبيل المثال عند إلقاء عملة متزنة 5 مرات حيث كل محاولة قد تكون صورة أو كتابة وذلك تحت فرض أن النجاح هو ظهور الصورة • هنا المحاولات المتكررة مستقلة كما أن احتمال النجاح ثابت من محاولة إلى أخرى ويساوي $p = \frac{1}{2}$ • ويجب ملاحظة أنه يمكن تعريف النجاح والفشل عكس ذلك تماماً، أي جعل ظهور الكتابة نجاح، ول هذه الحالة تتبدل قيمتي p, q •

عموماً يمكن القول أن تجربة ذي الحدين هي التجربة التي تحقق الشروط الآتية :-

- ١ - التجربة التي تتكون من n من المحاولات المتكررة .
- ب - نتيجة كل محاولة يمكن تصنيفها إلى نجاح أو فشل .
- ج - احتمال النجاح ، وهو p يبقى ثابت من محاولة إلى محاولة .
- د - المحاولات المتكررة مستقلة بعضها عن بعض .

ففي حالة إلقاء عملة 3 مرات ونحت فرض أن النجاح هو ظهور الصورة تعتبر عدد حالات النجاح متغير عشوائي يأخذ قيم صحيحة من 0 إلى 3 . النواتج الثمانية الممكنة وقيم x المقابلة معطاة في الجدول التالي .

النواتج	TTT	THT	TTH	HTT	THH	HTH	HHT	HHH
x	0	1	1	1	2	2	2	3

وحيث أن المحاولات مستقلة باحتمال نجاح ثابت ويساوي $\frac{1}{2}$ ، فإن :

$$P(HTH) = P(H)P(T)P(H) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}.$$

وبنفس الشكل، نحدد النواتج الأخرى باحتمال $\frac{1}{8}$. التوزيع الاحتمالي للمتغير العشوائي X

معطى في الجدول التالي :

x	0	1	2	3
$P(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

هذا ويمكن وضع صيغة للتوزيع الاحتمالي للمتغير العشوائي X على الشكل :-

$$f(x) = \frac{\binom{3}{x}}{8}, \quad x = 0, 1, 2, 3.$$

تعريف : عدد حالات النجاح X في n من المحاولات لتجربة ذي الحدين يسمى متغير عشوائي

يتبع ذي الحدين **binomial random variable** .

التوزيع الاحتمالي لمغير عشوائي X يتبع ذي الحدين يسمى توزيع ذي الحدين

binomial distribution وسوف نرمز له بالرمز $b(x; n, p)$ وذلك لأن قيمه تعتمد على

عدد المحاولات واحتمال النجاح في محاولة معطاه . وعلى ذلك عند إلقاء عملة 3 مرات فإن

التوزيع الاحتمالي للمتغير العشوائي X سوف يعاد كتابته كالآتي :

$$b(x; 3, \frac{1}{2}) = \frac{\binom{3}{x}}{8}, \quad x = 0, 1, 2, 3.$$

الآن دعنا نعمم المثال السابق للحصول على صيغة لـ $b(x; n, p)$. أي أننا نرغب في إيجاد صيغة تعطي احتمال الحصول على x حالات نجاح و $n-x$ حالات فشل في n من المحاولات لتجربة ذي الحدين . أولا نوجد احتمال الحصول على x حالات نجاح و $(n-x)$ حالات فشل في ترتيب معين . وحيث أن المحاولات مستقلة، فإنه يمكن ضرب كل الاحتمالات المقابلة للنواتج المختلفة . كل نجاح يحدث باحتمال p وكل فشل يحدث باحتمال $q = 1 - p$. وعلى ذلك، الاحتمال للترتيب المعين هو $p^x q^{n-x}$ وبالطبع ترتيب الضرب هنا غير مهم . الآن لا بد من حساب العدد الكلي من نقاط العينة في التجربة (الترتيبات) والتي تعطي x حالات نجاح و $(n-x)$ حالات فشل . هذا العدد يساوي عدد الطرق لتبديل n من العناصر حيث x منهم من نوع (نجاح) و $(n-x)$ من نوع آخر (فشل) . وهذا يحدث بطرق عددها $\binom{n}{x}$. وحيث أن الترتيبات التي عددها $\binom{n}{x}$ متتالية في إمكانية الحدوث وتحدث باحتمال متساوي ، فإنه يمكن الحصول على توزيع ذي الحدين بجمع الاحتمالات لكل الترتيبات أو ببساطة ضرب $p^x q^{n-x}$ في $\binom{n}{x}$. وعلى ذلك فإن الصيغة العامة لتوزيع ذي الحدين سوف تكون على الشكل :

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, \dots, n.$$

ويجب أن نتذكر أنه في مثال إلقاء عملة 3 مرات حيث $p = \frac{1}{2}$ ، $n = 3$ فإن :

$$b(x; 3, \frac{1}{2}) = \binom{3}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{3-x} = \frac{\binom{3}{x}}{8}$$

والتي تتفق مع الإجابة التي سبق أن حصلنا عليها .

في الحقيقة فإن التوزيع الاحتمالي لذي الحدين أشتق اسمه من أن $(n+1)$ من الحدود في

مفكوك ذي الحدين $(p+q)^n$ تقابل لهم $b(x; n, p)$ ، حيث $x = 0, 1, 2, \dots, n$ ، أي أن :

$$(p+q)^n = \binom{n}{0} q^n + \binom{n}{1} p q^{n-1} + \binom{n}{2} p^2 q^{n-2} + \dots + \binom{n}{n} p^n \\ = b(0; n, p) + b(1; n, p) + \dots + b(n; n, p).$$

وحيث أن $p+q = 1$ فإننا نجد أن $\sum_{x=0}^n b(x; n, p) = 1$ وهو الشرط الذي لا بد من تحققه

لأي توزيع احتمالي . في الحقيقة سوف يكون الاهتمام بإيجاد $P(X \leq r)$ أو

$P(a \leq X \leq b)$. لحسن الحظ فإن $\sum_{x=0}^r b(x; n, p)$ معطاة في جدول ذي الحدين في

ملحق (١) حيث $n=5, 10, 15, 20, 25$.

مثال (٦-٣) إذا أُلقيت زهرة نرد متزنة 6 مرات • أوجد احتمال ظهور الرقم 5 أربع مرات بالضبط •

الحل • هنا يعتبر ظهور الرقم 5 نجاح وعلى ذلك احتمال النجاح في كل محاولة من المحاولات الستة المستقلة هو $\frac{1}{6}$ • وعلى ذلك احتمال (الفشل) عدم ظهور الرقم خمسة هو $\frac{5}{6}$ • بفرض أن X تمثل عدد مرات ظهور الرقم 5 ولها دالة كثافة الاحتمال :

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, \dots, n.$$

وعلى ذلك الاحتمال المطلوب هو :

$$b(4; 6, \frac{1}{6}) = \binom{6}{4} \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^2$$

$$= \frac{6!}{4!2!} \cdot \frac{5^2}{6^6} = 0.00803755.$$

مثال (٦-٤) صندوق به 10 ثمرات منها 3 تالفة، اختيرت منه ثمرتين • أحسب احتمال أن تكون واحدة تالفة (السحب بإرجاع) •

الحل • بفرض أن X تمثل عدد الثمار التالفة ولها دالة كثافة الاحتمال :

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, \dots, n.$$

وعلى ذلك الاحتمال المطلوب هو :

$$b(1; 2, 0.3) = \binom{2}{1} (0.3)^1 (0.7)^1$$

$$= \frac{2!}{1!1!} \cdot (0.3)^1 (0.7)^1 = 0.42.$$

مثال (٦-٥) احتمال أن يشفى مريض من مرض نادر في الدم هو 0.2، فإذا كان معروف أن 15 شخص عندهم هذا المرض ما هو احتمال أن (أ) يشفى 9 على الأقل من المرض (ب) يشفى من 4 إلى 8 من هذا المرض (جـ) يشفى على الأكثر أثنين من هذا المرض؟

الحل • (أ) بفرض أن X تمثل عدد المرضى الذين سوف يشقوا من هذا المرض فإن :

$$P(x \geq 9) = 1 - P(x < 9) = 1 - P(x \leq 8)$$

$$= 1 - \sum_{x=0}^8 b(x; 15, 0.2) = 1 - 0.999 = 0.001.$$

$$P(4 \leq X \leq 8) = \sum_{x=0}^8 b(x; 15, 0.2) - \sum_{x=0}^3 b(x; 15, 0.2) \quad (\text{ب})$$

$$= 0.999 - 0.648 = 0.351.$$

$$P(X \leq 2) = \sum_{x=0}^2 b(x; 15, 0.2) = 0.3980. \quad (\text{جـ})$$

نظرية (١-٦) الوسط الحسابي والتباين لتوزيع ذي الحدين $b(x; n, p)$ هما :

$$\mu = pq, \quad \sigma^2 = npq$$

البرهان : بفرض أن نتيجة المحاولة رقم j يمثل بالمعبر العشوائي I_j والذي يسمى (مضرب برنولي) والذي يأخذ القيمة 0 أو 1 باحتمالات p, q على التوالي، أي أن $I_j = 0$ تدل على فشل و $I_j = 1$ تدل على نجاح. وعلى ذلك، في تجربة ذي الحدين يمكن كتابة عدد حالات النجاح كحاصل جمع لعدد n من مضربات برنولي المستقلة. وعلى ذلك :

$$X = I_1 + I_2 + \dots + I_n.$$

الوسط الحسابي لأي I_j هو $E(I_j) = 0 \cdot q + 1 \cdot p = p$ وعلى ذلك، وباستخدام نظرية (٦-٤) فإن الوسط الحسابي لتوزيع ذي الحدين هو :

$$\begin{aligned} \mu &= E(X) = E(I_1) + E(I_2) + \dots + E(I_n) \\ &= p + p + \dots + p = np. \end{aligned}$$

وذلك لأن عدد الحدود يساوي n .

التباين لأي I_j هو :

$$\begin{aligned} \sigma_{I_j}^2 &= E(I_j - p)^2 = E(I_j^2) - p^2 \\ &= (0)^2 q + (1)^2 p - p^2 = p(1 - p) = pq. \end{aligned}$$

وعلى ذلك، وباستخدام نظرية (٨-٤) فإن التباين لتوزيع ذي الحدين هو :

$$\begin{aligned} \sigma_X^2 &= \sigma_{I_1}^2 + \sigma_{I_2}^2 + \dots + \sigma_{I_n}^2 \\ &= pq + pq + \dots + pq = npq. \end{aligned}$$

مثال (٦-٦) أوجد الوسط الحسابي والتباين لتوزيع ذي الحدين في مثال (٥-٦) .

الحل . حيث $p = 0.2$, $n = 15$ وباستخدام نظرية (١-٦) فإن :

$$\mu = (15)(0.2) = 3, \quad \sigma^2 = (15)(0.2)(0.8) = 2.4.$$

تجربة ذي الحدين تسمى تجربة متعددة الحدود multinomial experiment إذا

كانت كل محاولة لها k من النواتج حيث $k \geq 2$. عموماً إذا كانت محاولة معطاة k من النواتج الممكنة E_1, E_2, \dots, E_k باحتمالات p_1, p_2, \dots, p_k ، فإن التوزيع المتعدد الحدود سوف يعطى الاحتمال أن E_1 تحدث x_1 من المرات وأن E_2 تحدث x_2 من المرات و... و E_k تحدث x_k من المرات في n من المحاولات المستقلة، حيث $x_1 + x_2 + \dots + x_k = n$ سوف

نرمز للتوزيع المتعدد الحدود بالرمز $f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n)$ من الواضح أن $p_1 + p_2 + \dots + p_k = 1$ لأن نتيجة كل محاولة لابد أن تكون واحدة من k من النواتج الممكنة. لاشتقاق الصيغة العامة سوف نتبع نفس الأسلوب المستخدم في توزيع ذي الحدين. وحيث أن المحاولات مستقلة وأنه في ترتيب معين نحصل على x_1 نواتج من E_1 و x_2 من E_2 و \dots و x_k من E_k باحتمال $p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$ العدد الكلي من الترتيبات والذي يؤدي إلى نفس النتائج في n من المحاولات المستقلة يساوي عدد الطرق لتبديل n من العناصر حيث x_1 منها تنتمي إلى E_1 و x_2 تنتمي إلى E_2 و \dots و x_k تنتمي إلى E_k والذي يحدث بطرق عددها $\frac{n!}{x_1! x_2! \dots x_k!}$ وحيث أن الترتيبات (نقاط العينة) التي عددها $\frac{n!}{x_1! x_2! \dots x_k!}$ متساوية في إمكانية الحدوث وتحدث باحتمال متساوي، فإنه يمكن الحصول على التوزيع المتعدد الحدود بضرب الاحتمال لترتيب معين في $\frac{n!}{x_1! x_2! \dots x_k!}$ وعلى ذلك فإن التوزيع متعدد الحدود يكون على الصورة :

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n)$$

$$= \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k},$$

حيث $\sum_{i=1}^k p_i = 1, \sum_{i=1}^k x_i = n$ وقد اشتق التوزيع المتعدد الحدود اسمه من أن الحدود في

المكسوك $(p_1 + p_2 + \dots + p_k)^n$ تقابل كل القيم الممكنة للبالدة

$$f(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k, n)$$

مثال (٦-٧) ألقيت زهرة نرد متزنة 11 مرة أوجد احتمال الحصول على رقم واحد مرة ،

ورقم 2 مرتين ورقم 3 خمس مرات ورقم أربعة مرتين ورقم 6 مرة واحدة.

الحل . احتمال ظهور أي رقم عند إلقاء زهرة نرد متزنة مرة واحدة متساوي ويساوي $\frac{1}{6}$ وعلى

ذلك فإن الاحتمال المطلوب هو :

$$\begin{aligned}
 & f(1,2,5,2,1; \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, 11) \\
 &= \frac{11!}{1!2!5!2!1!} \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^5 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^1 \\
 &= \frac{11!}{1!2!5!2!1!} \cdot \frac{1}{6^{11}} = 0.000229219.
 \end{aligned}$$

(٣-٦) التوزيع الهندسي الزائدى Hypergeometric Distribution

بفرض أن مجتمع يتكون من عدد محدود من الوحدات، وليكن N ، وأن هناك k مسن الوحدات من النوع A (نجاح) والوحدات الباقية من نوع B (فشل)، و بفرض أن عينة عشوائية من الحجم n اختيرت من هذا المجتمع بدون إرجاع. بفرض أن X تمثل عدد الوحدات من نوع A التي تظهر في العينة. اهتمامنا سوف يكون في إيجاد $P(X=x)$. التجربة السابقة تسمى تجربة الهندسي الزائدى **hypergeometric experiment**. تحقق تجربة الهندسي الزائدى الشروط التالية:

- أ- عينة عشوائية من الحجم n تختار من مجتمع يتوى على N من الوحدات.
 - ب- في المجتمع الذي حجمه N فإن k من الوحدات تصنف نجاح و $N - k$ تصنف فشل.
- تعريف: عدد حالات النجاح في تجربة الهندسي الزائدى يسمى متغير عشوائي يتبع الهندسي الزائدى.

التوزيع الاحتمالي لمتغير عشوائي يتبع التوزيع الهندسي الزائدى يسمى التوزيع الهندسي الزائدى ويمثل بالرمز $h(x; N, n, k)$ وذلك لأن عدد حالات النجاح x تعتمد على k الموجودة في الفئة N ، حيث يختار من N وحدات عددها n . في المثال (٦-٤) استخدمنا توزيع ذي الحدين إذا كان السحب بإرجاع (المحاولات مستقلة). الآن بفرض أن السحب بدون إرجاع (المحاولات غير مستقلة). في هذه الحالة سوف يكون هناك $\binom{3}{1}$ طريقة لاختيار ثمرة تالفة ولكل واحدة من هذه الطرق يوجد $\binom{7}{1}$ طريقة لاختيار ثمرة سليمة. وعلى ذلك عند اختيار ثمرتين من الصندوق بدون إرجاع فإن عدد الطرق الكلية للحصول على ثمرة تالفة و ثمرة سليمة هو $\binom{3}{1} \binom{7}{1}$. العدد الكلى من الطرق لاختيار ثمرتين من الصندوق المختوى على 10 ثمرات هو $\binom{10}{2}$. وعلى ذلك احتمال الحصول على ثمرة تالفة و ثمرة سليمة عند اختيار عينة من الحجم $n=2$ من الصندوق المختوى على 10 ثمرات هو:

$$\frac{\binom{3}{1}\binom{7}{1}}{\binom{10}{2}} = \frac{7}{15}.$$

المثال السابق يوضح ما يسمى بتجربة الهندسي الزائدى .

مثال (٦-٨) يراد اختيار لجنة من ثلاثة أشخاص من بين 4 سيدات و 5 رجال المطلوب إيجاد التوزيع الاحتمالي لعدد السيدات في اللجنة المختارة .

الحل . بفرض أن الطير العشوائي X يمثل عدد السيدات في اللجنة المختارة . الشروط لتجربة الهندسي الزائد متوفرة وعلى ذلك :

$$P(X=0) = h(0;9,3,4) = \frac{\binom{4}{0}\binom{5}{3}}{\binom{9}{3}} = \frac{10}{84},$$

$$P(X=1) = h(1;9,3,4) = \frac{\binom{4}{1}\binom{5}{2}}{\binom{9}{3}} = \frac{40}{84},$$

$$P(X=2) = h(2;9,3,4) = \frac{\binom{4}{2}\binom{5}{1}}{\binom{9}{3}} = \frac{30}{84},$$

$$P(X=3) = h(3;9,3,4) = \frac{\binom{4}{3}\binom{5}{0}}{\binom{9}{3}} = \frac{4}{84}.$$

ويمكن تمثيل التوزيع الهندسي الزائدى بالجدول التالي :

x	0	1	2	3
$P(X=x)$	$\frac{10}{84}$	$\frac{40}{84}$	$\frac{30}{84}$	$\frac{4}{84}$

هذا ويمكن وضع صيغة لإيجاد التوزيع الاحتمالي للمثال السابق على الشكل :

$$P(X=x) = h(x;9,3,4) = \frac{\binom{4}{x}\binom{5}{3-x}}{\binom{9}{3}}, x=0,1,2,3.$$

الآن يمكن تعميم صيغة التوزيع الاحتمالي في المثال السابق وذلك للحصول على صيغة

للدالة $h(x;N,n,k)$. العدد الكلى للعينات من الحجم n المختارة من N من الوحدات هو

$\binom{N}{n}$. هذه العينات يفترض أنها متساوية في إمكانية الحدوث . يوجد $\binom{k}{x}$ طريقة لاختيار x

من k حالات النجاح ، ولكل طريقة من هذه الطرق يمكن اختيار $(n-x)$ من حالات الفشل بطرق عددها $\cdot \binom{N-k}{n-x}$. وعلى ذلك العدد الكلى من العينات المرغوب فيها من $\binom{N}{n}$ عينة هو $\cdot \binom{k}{x} \binom{N-k}{n-x}$. وعلى ذلك فإن التوزيع الهندسي الزائدى يمكن الحصول عليه كالآتي :

$$h(x; N, n, k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}, x = 0, 1, 2, \dots, n.$$

مثال (٩-٦) اختبرت عينة عشوائية من الحجم $n=6$ من صندوق يحوى على 5 كرات حمراء و 4 كرات سوداء . ما هو احتمال ظهور ثلاث كرات حمراء في العينة المختارة .
الحل . باستخدام التوزيع الهندسي الزائدى حيث $x = 3, k = 5, n = 6, N = 9$ ، فإن :

$$h(3; 9, 6, 5) = \frac{\binom{5}{3} \binom{4}{3}}{\binom{9}{6}} = \frac{40}{84}.$$

نظرية (٦-٢) الوسط الحسابي والتباين للتوزيع الهندسي الزائدى هما :

$$\mu = \frac{nk}{N}, \sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N} \left(1 - \frac{k}{N}\right).$$

إذا كانت n صغيرة بالنسبة إلى N فإنه يمكن استخدام توزيع ذي الحدين كتقريب للتوزيع الهندسي الزائدى حيث $p = \frac{k}{N}$. وعلى ذلك يمكن تقريب الوسط الحسابي والتباين بالصيغة :

$$\mu = np = \frac{nk}{N}, \sigma^2 = npq = n \cdot \frac{k}{N} \left(1 - \frac{k}{N}\right).$$

بمقارنة الصيغتين السابقتين بالصيغتين في نظرية (٦-٢) فإننا نجد أن الوسط الحسابي هو نفسه بينما التباين يختلف بمعامل تصحيح $\frac{N-n}{N-1}$ وهذا يمكن إهماله عندما تكون n صغيرة بالنسبة إلى N .

مثال (٦-١٠) في سترال وجد أنه من بين 4000 تليفون تم تركيبهم في منطقة حديثة يوجد 3000 منهم يختلف لونهم عن اللون الأسود . تحدث 5 أشخاص عشوائيا، فما هو احتمال أن 2 منهم بالضبط سوف يتحدثون من تليفون لونه أسود .

الحل . وحيث أن المجتمع حجمه $N=4000$ وهذا الحجم كبير نسبيا بالنسبة لحجم العينة $n=5$ فإننا سوف نقرب الاحتمال المطلوب باستخدام توزيع ذي الحدين . احتمال أن شخص يتحدث من تليفون أسود هو 0.25 وعلى ذلك احتمال أن 2 بالضبط يتصلون من تليفون أسود هو :

$$h(2;4000,5,1000) \approx b(2;5,0.25)$$

$$= \sum_{x=0}^2 b(x;5,0.25) - \sum_{x=0}^1 b(x;5,0.25) = 0.896 - 0.633 = 0.263.$$

يمكن تعميم التوزيع الهندسي الزائدى ليعامل للحالة التي يكون فيها المجتمع مقسم الى k من الخلايا A_1, A_2, \dots, A_k حيث a_1 وحدة في الخلية A_1 و a_2 وحدة في الخلية A_2 و \dots و a_k وحدة في الخلية A_k . الآن سوف يكون اهتمامنا في إيجاد صيغة لاحتمال أن العينة العشوائية من الحجم n سوف تعطى x_1 عنصر من A_1 و x_2 عنصر من A_2 و \dots و x_k عنصر من A_k والتي سوف تمثل الاحتمال بالصيغة :

$$f(x_1, x_2, \dots, x_k; a_1, a_2, \dots, a_k, N, n).$$

لإيجاد صيغة عامة فإننا نعرف أن العدد الكلى من العينات من الحجم n التي يمكن اختيارها من مجتمع حجمه N هو $\binom{N}{n}$. وعلى ذلك هناك $\binom{a_1}{x_1}$ طريقة لاختيار x_1 من الوحدات الموجودة في A_1 ولكل واحدة من هذه الطرق فإننا يمكن اختيار x_2 من الوحدات الموجودة في A_2 بطرق عددها $\binom{a_2}{x_2}$. وعلى ذلك يمكن اختيار x_1 من الوحدات الموجودة في A_1 واختيار x_2 من الوحدات الموجودة في A_2 بطرق عددها $\binom{a_1}{x_1} \binom{a_2}{x_2}$ وبالاتسار في هذا الأسلوب يمكن اختيار كل الوحدات التي عددها n والتي تحتوى على x_1 من الوحدات الموجودة في A_1 و x_2 من الوحدات الموجودة في A_2 و \dots و x_k من الوحدات الموجودة في A_k بطرق عددها $\binom{a_1}{x_1} \binom{a_2}{x_2} \dots \binom{a_k}{x_k}$. وعلى ذلك فإن التوزيع الاحتمالي المطلوب سوف يكون :

$$f(x_1, x_2, \dots, x_k; a_1, a_2, \dots, a_k, N, n) = \frac{\binom{a_1}{x_1} \binom{a_2}{x_2} \dots \binom{a_k}{x_k}}{\binom{N}{n}}.$$

$$\text{حيث } \sum_{i=1}^k x_i = n, \sum_{i=1}^k a_i = N$$

مثال (٦-١١) يحتوى صندوق على 3 كرات حمراء و 4 كرات بيضاء و 5 كرات سوداء . يرغب شخص في اختيار 4 كرات من الصندوق، ما هو الاحتمال أن يختار كرة حمراء و 2 كرة بيضاء و واحدة سوداء ؟

الحل . باستخدام الصيغة العامة للتوزيع الهندسي الزائدى حيث

$$: \text{فإن } a_1 = 3, x_1 = 1, a_2 = 4, x_2 = 2, a_3 = 5, x_3 = 1, N = 12, n = 4$$

$$f(1,2,1;3,4,5,12,4) = \frac{\binom{3}{1} \binom{4}{2} \binom{5}{1}}{\binom{12}{4}} = \frac{90}{495}.$$

(٤-٦) توزيع بواسون Poisson Distribution

أن التجارب التي تعطينا عدد حالات النجاح والتي تحدث في فترة زمنية معينة أو في منطقة محددة، تسمى تجارب بواسون poisson experiment . الفترة الزمنية المعنية قد تكون دقيقة، يوم، أسبوع، شهر أو حتى سنة . وعلى ذلك تجربة بواسون قد تنتج مشاهدات لمغير عشوائي X يمثل عدد المكالمات التليفونية في الساعة والمستقبل من مكعب، أو عدد الأيام في السنة والتي تغلق فيها بعض المدارس بسبب الصقيع في بلد ما . المنطقة المحددة يمكن أن تكون خط الأعداد الحقيقية، مساحة، حجم أو ربما قطعة من المعدن . في هذه الحالة X يمكن أن تمثل عدد الفئران في فدان من القمح، عدد البكتريا في لتر من الماء النقي، عدد الأخطاء في صفحة من قاموس . تجارب بواسون يجب أن تحقق الشروط التالية :

أ- متوسط عدد حالات النجاح، μ ، والتي تحدث في الفترة الزمنية المعطاة أو في المنطقة المحددة معلوم .

ب- احتمال وقوع حالة نجاح واحدة في فترة قصيرة أو منطقة صغيرة يتناسب مع طول هذه الفترة أو حجم هذه المنطقة ولا يعتمد على عدد حالات النجاح التي تحدث خارج هذه الفترة أو المنطقة .

ج- احتمال وقوع أكثر من حالة نجاح في الفترة القصيرة أو المنطقة الصغيرة مهمل .

تعريف : عدد حالات النجاح X في تجربة بواسون يسمى متغير عشوائي يتبع بواسون .

التوزيع الاحتمالي لمغير عشوائي X يتبع توزيع بواسون يمثل بالصيغة $p(x; \mu)$ وذلك لأن قيمته تعتمد فقط على μ . متوسط عدد حالات النجاح التي تحدث في الفترة المعنية أو المنطقة المحددة تساوي التباين . اشتقاق توزيع بواسون خارج نطاق هذا الكتاب . التوزيع الاحتمالي لمغير عشوائي يتبع بواسون هو :

$$p(x; \mu) = \frac{e^{-\mu} \mu^x}{x!}, x = 0, 1, 2, \dots,$$

حيث μ متوسط عدد حالات النجاح التي تحدث في الفترة المعطاة أو المنطقة الخاصة و $e=2.71828...$

الجدول في ملحق (٢) يحتوي على حاصل جمع احتمالات بواسون أي $\sum_{x=0}^{\infty} p(x; \mu)$

لقيم محددة من μ تتراوح بين 0.1 إلى 20 . سوف نشرح طريقة عمل هذه الجداول بالمثلثين التاليين .

مثال (٦-١٢) نُفْطَل مَآكِنَة لِتَصْنِيعِ الْخُلُوفِ فِي الْمَوَسَطِ خَمْسَ مَرَاتٍ فِي الْأَسْبُوعِ مَا هُوَ الْإِحْتِمَالُ أَنَّ تَعَطَّلَ الْمَآكِنَة ثَلَاثَ مَرَاتٍ خِلَالِ اسْبُوعٍ .

الحل . باستخدام توزيع بواسون حيث $\mu = 5$, $x = 3$ ومن جدول بواسون في الملحق (٢) فإن :

$$p(3;5) = \frac{e^{-5}5^3}{3!} = \sum_{x=0}^3 p(x;5) - \sum_{x=0}^2 p(x;5)$$

$$0.265 - 0.125 = 0.14.$$

مثال (٦-١٣) إذا كان متوسط عدد الفترات في فدان من القمح هو $\mu = 3$. أوجد احتمال وجود على الأقل 12 قارا في فدان معطى .

الحل . نفرض أن X تمثل عدد الفترات في فدان من القمح وباستخدام جداول بواسون في ملحق (٢) فإن :

$$P(X \geq 12) = 1 - P(X < 12)$$

$$= 1 - \sum_{x=0}^{11} p(x;3)$$

$$= 1 - P(X \leq 11) = 1 - 1.000 = 0.000.$$

المبرج التكراري لكل من توزيع ذي الحدين وتوزيع بواسون تقريبا هم نفس الشكل عندما تكون n كبيرة و p تقرب من الصفر وعلى ذلك إذا تحقق هذا الشرط، فإن توزيع بواسون بمتوسط $\mu = np$ ، يمكن استخدامه كتقريب لإيجاد احتمالات توزيع ذي الحدين . إذا اقتربت p من الواحد يمكن تغيير ما قمنا بتعريفه لنجاح إلى فشل والعكس الفشل إلى نجاح . وعلى ذلك تنفر p إلى قيمة قريبة من الصفر

مثال (٦-١٤) يحتوى كتاب على 1000 صفحة يوجد ما 400 خطأ موزعة على صفحات الكتاب ما هو احتمال الحصول على ستة أخطاء في صفحة مختارة ؟

الحل . المتغير العشوائي X يمثل عدد الأخطاء في إحدى الصفحات وهو يتبع توزيع ذي الحدين بمعلمتي p , n حيث $p = \frac{1}{1000}$ و $n=400$ وحيث أن n كبيرة و p صغيرة فإن توزيع X يؤول إلى توزيع بواسون بمعلمة $\mu = np = (400) \times (\frac{1}{1000}) = 0.4$ وباستخدام جداول بواسون في الملحق (٢) حيث $\mu = 0.4$ فإن :

$$p(x = 6) = b(x; 400, 0.001) \approx p(6; 0.4),$$

$$i(6;0.4) = \frac{e^{-0.4} 0.4^6}{6!} = \sum_{x=0}^6 p(x;0.4) - \sum_{x=0}^5 p(x;0.4) = 1.000 - 1.000 = 0.000.$$

مثال (٦-١) يفرض أنه في المتوسط يوجد شخص من بين 1000 شخص يدخن السجائر في مدينة ما ، أوجد احتمال أنه في عينة عشوائية من 8000 شخص نحصل على 6 أشخاص يدخنون .

الحل . في هذا المثال يكون لدينا تجربة ذي الحدين حيث $n=8000$ ، $p=0.001$ وحيث أن p تقرب كثيرا من الصفر وأن n كبيرة بل درجة كالية، فإنه يمكن استخدام توزيع بواسون حيث $\mu = (8000)(0.001) = 8$ ، وعلى ذلك إذا كانت X تمثل عدد المدخنين، فإن :

$$\begin{aligned} p(X = 6) &= b(x; 8000, 0.001) \\ &\approx p(6; 8) = \sum_{x=0}^6 p(x; 8) - \sum_{x=0}^5 p(x; 8) \\ &= 0.313 - 0.191 = 0.122. \end{aligned}$$

(٦-٥) توزيع ذي الحدين السالب Negative Binomial Distribution

يفرض أن تجربة ما لها نفس الخصائص التي سبق أن ذكرناها لتوزيع ذي الحدين، ولكن مع تكرار المحاولات حتى يمكن الحصول على عدد ثابت من حالات النجاح . وعلى ذلك، بدلا من إيجاد احتمال الحصول على x نجاح في n من المحاولات، فإن الاهتمام سوف يكون في إيجاد احتمال أن النجاح رقم x سوف يحدث في المحاولة رقم x . التجارب من هذا النوع تسمى تجارب ذي الحدين السالب **negative binomial experiment** . يفرض أن لاعب كرة السلة يتبع في التصويب نحو الهدف في 80% من المحاولات . المطلوب إيجاد احتمال أن يتبع في تصويب الهدف الخامس في المحاولة رقم 8 . سوف نرمز للنجاح في التصويب بالرمز F' ولنرمز للفشل في التصويب بالرمز F ، وعلى ذلك لترتيب مطلوب سوف نحصل على $F'F'F'F'F'$ والذي يحدث باحتمال:

$(0.2)^5 (0.8)^3 = (0.8)(0.8)(0.8)(0.2)(0.2)(0.2)(0.8)(0.8)(0.8)$. ويمكن حصر كل الترتيبات بإعادة ترتيب حالات النجاح والفشل ما عدا المحاولة الأخيرة، والتي لا بد أن تكون النجاح رقم خمسة . العدد الكلي من الترتيبات الممكنة يساوي عدد الطرق لتبديل 7 عناصر منها 4 من نوع نجاح و 3 من نوع فشل . هذا العدد الكلي من الترتيبات الممكنة يحدث بطرق متنافية عددها $\left(\frac{7}{4}\right)$. وحيث أن X تمثل عدد المحاولات اللازمة للحصول على 5 أهداف، فإن :

$$P(X = 8) = \left(\frac{7}{4}\right)(0.8)^5 (0.2)^3 = 0.0917504.$$

تعريف : العدد X من المحاولات والذي ينتج k حالات نجاح في تجربة ذي الحدين السالب يسمى
مطير عشوائي يتبع ذي الحدين السالب .

التوزيع الاحتمالي لمطير عشوائي يتبع ذي الحدين السالب يسمى توزيع ذي الحدين السالب، وسوف نمثله بالصيغة $b^*(x; k, p)$ حيث أن قيمه تعتمد على عدد حالات النجاح المطلوبة واحتمال النجاح في المحاولة المعطاه . لإيجاد الصيغة العامة سوف نوجد أولاً احتمال الحصول على

نجاح في المحاولة رقم x والتي تكون مسبوقة بعدد $(k-1)$ من حالات النجاح و $(x-k)$ من حالات الفشل وذلك في ترتيب خاص . وحيث أن محاولات مستقلة، فإنه يمكن ضرب كل الاحتمالات المقابلة لكل نتيجة مطلوبة . كل حالة نجاح تحدث باحتمال p وكل حالة فشل تحدث باحتمال $q=1-p$. وعلى ذلك، الاحتمال أن الترتيب المطلوب ينتهي بحالة نجاح هو $p^{k-1}q^{x-k}p = p^kq^{x-k}$. العدد الكلي من نقاط العينة في التجربة والذي ينتهي بحالة نجاح، بعد $k-1$ من حالات النجاح و $x-k$ من حالات الفشل في أي ترتيب، يساوي عدد الطرق لنيل $x-1$ من المحاولات منهم $k-1$ حالات نجاح و $x-k$ حالات فشل والتي تساوي $\binom{x-1}{k-1}$. كل الترتيبات المطلوبة تكون متافيه في إمكانية الحدوث وتحدث باحتمال متساوي وهو p^kq^{n-k} ، وعلى ذلك يمكن الحصول على الاحتمال المطلوب بضرب p^kq^{x-k} في $\binom{x-1}{k-1}$ ، وعلى ذلك توزيع ذي الحدين السالب يأخذ الصيغة التالية :

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}, \quad x = k, k+1, k+2, \dots$$

مثال (٦-١٦) : إذا كان احتمال ولادة ذكر في أي ولادة غير لها سيدة هو $\frac{1}{2}$ أوجد احتمال أن تضع ذكراً بعد أربع ولادات .

الحل . باستخدام توزيع ذي الحدين السالب حيث $k=2, x=4, p=\frac{1}{2}$ فإن :

$$\begin{aligned} b^*(4; 2, 0.5) &= \binom{3}{1} 0.5^2 0.5^2 \\ &= \frac{3!}{1!2!} \cdot \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = \frac{3}{16} \end{aligned}$$

في الحقيقة أشتق اسم توزيع ذي الحدين السالب من أن كل حد في المكوك
 $\bullet x = k, k+1, k+2, \dots$ حيث $b^*(x; k, p)$ يقابل قيمة من قيم $p^k(1-q)^{n-k}$
 عندما $n=1$ فإننا نحصل على الحالة الخاصة من توزيع ذي الحدين السالب، أي نحصل على
 التوزيع الاحتمالي لعدد المحاولات المطلوبة للحصول على حالة نجاح واحدة. توزيع ذي الحدين
 السالب سوف يتحول إلى الشكل :

$$b^*(x; 1, p) = pq^{x-1}, x = 1, 2, 3, \dots$$

والذي يسمى التوزيع الهندسي وسوف نرمز له بالرمز $g(x; p)$
 مثال (٦-١٧) أوجد (أ) التوزيع الاحتمالي لعدد المحاولات اللازمة للحصول على صورة
 واحدة وذلك عند إلقاء عملة متزنة (ب) احتمال الحصول على صورة في المحاولة الرابعة.

الحل . باستخدام التوزيع الهندسي حيث $p = \frac{1}{2}$ نحصل على :

$$g(x; 0.5) = (0.5)(0.5)^{x-1}, x = 1, 2, 3, \dots \quad (أ)$$

$$g(4; \frac{1}{2}) = (0.5)(0.5)^3 = \frac{1}{16}. \quad (ب)$$

Normal Distribution (٦-٦) التوزيع الطبيعي

يعبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية المتصلة في علم الإحصاء، حيث
 يصف كثيرا من المجتمعات الموجودة في الطبيعة، الصناعة، الأبحاث. دالة كثافة الاحتمال للتوزيع
 الطبيعي هي:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

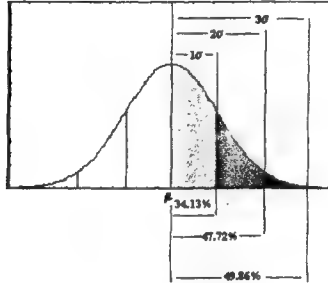
على الفترة $-\infty < x < \infty$ حيث $e = 2.71828 \dots$ و $\pi = 3.14159 \dots$ و μ و σ هما
 الوسط الحسابي (المتوسط) والانحراف المعياري على التوالي. بيان $f(x)$ في شكل (٦-٢) على
 شكل ناقوس متمائل حول العمود المقام على الوسط الحسابي. ينطبق الوسط الحسابي على
 الوسيط وأيضا على المتوسط. يتقارب طرفا منحنى التوزيع الطبيعي من الصفر عندما
 $\bullet x \rightarrow -\infty$ أو $x \rightarrow \infty$

لاي توزيع طبيعي فإن :-

(أ) المساحة الواقعة تحت المنحنى الطبيعي بين النقطتين μ و $(\mu + \sigma)$ تمثل 34.13% من

المساحة الكلية كما يتضح من شكل (٦-٢).

- (ب) المساحة الواقعة تحت المنحنى الطبيعي بين النقطتين μ و $(\mu+2\sigma)$ تمثل 47.72 من المساحة الكلية كما يتضح من شكل (٢-٦).
- (جـ) المساحة الواقعة تحت المنحنى الطبيعي بين النقطتين μ و $(\mu+3\sigma)$ تمثل 49.86 % من المساحة الكلية كما يتضح من شكل (٢-٦).



شكل (٢-٦)

وبما أن المنحنى الطبيعي متماثل فإن القيم السابقة تتحقق عند طرح الانحراف المعياري من المتوسط .

(١-٦-٦) التوزيع الطبيعي القياسي :

Standard Normal Distribution

إذا كان التوزيع الطبيعي متوسطه صفر وتباينه الواحد الصحيح فإنه يسمى التوزيع الطبيعي القياسي . بفرض أن Z ترمز لمتغير عشوائي متصل له توزيع طبيعي قياسي فإن دالة الكثافة الاحتمالية لذلك المتغير تأخذ الشكل الآتي :

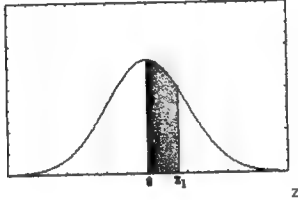
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty \leq z \leq \infty.$$

إذا كان z_1 عدد حقيقي موجب فإن الاحتمال $(0 < Z < z_1)$ يساوي المساحة المظللة في شكل (٣-٦)، ويمكن الحصول عليها من جدول التوزيع الطبيعي القياسي في ملحق (٣) . المساحات الواقعة تحت المنحنى الطبيعي القياسي غير معطاة في الجدول لقيم Z السالبة ولكن يمكن حسابهم باستخدام خاصية التماثل للمنحنى الطبيعي . المساحة الواقعة تحت المنحنى الطبيعي

القياسي بين $z = 0$ و $z = -z_1$ تساوي المساحة الواقعة بين $z = z_1$ و $z = 0$ أي أن :

$$P(-z_1 < Z < 0) = P(0 < Z < z_1)$$

معظم المساحات تحت المنحنى الطبيعي القياسي تقع في الفترة (3,-3) ونادرا ما نجد قيم تقع خارج هذه الفترة .



شكل (٦-٣)

مثال (٦-١٨) إذا كان Z متغير عشوائي يتبع التوزيع الطبيعي القياسي احسب الاحتمالات الآتية مع توضيح ذلك بإيادنا .

(أ) $P(0 < Z < 1.05)$ (ب) $P(-1.06 < Z < 1.06)$

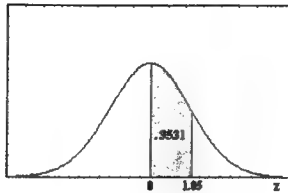
(جـ) $P(-0.47 < Z < 0.95)$ (د) $P(1.6 < Z < 2)$

(هـ) $P(Z > 2.02)$ (و) $P(Z < -0.45)$ (ز) $P(Z < 1.07)$

الحل . (أ) لإيجاد قيمة الاحتمال $P(0 < Z < 1.05)$ نبحث في العمود الأول على الشمال من جدول التوزيع الطبيعي القياسي في ملحق (٣) عن القيمة 1.0 ثم نتحرك أمام هذه القيمة أفقيا حتى نصل إلى العمود الذي رأس عنوانه الرقم 0.05 فتكون هي المساحة المطلوبة أي أن :

$$P(0 < Z < 1.05) = 0.3531.$$

والتي تمثل المساحة المظلة في شكل (٦-٤) .



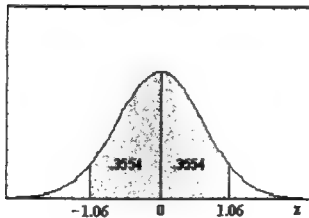
شكل (٤-٦)

(ب) الاحتمال المطلوب هو $P(-1.06 < Z < 1.06)$ وهو يساوي المساحة المظللة في شكل (٤-٦)

• (٥-٦)

ونظرا لتمثيل المنحنى الطبيعي فإن :

$$\begin{aligned} P(-1.06 < Z < 1.06) &= P(-1.06 < Z < 0) + P(0 < Z < 1.06) \\ &= 2P(0 < Z < 1.06) = 2(0.3554) = 0.7108. \end{aligned}$$

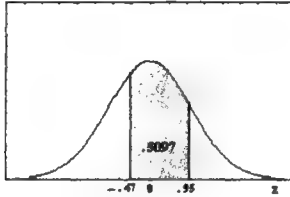


شكل (٥-٦)

(جـ) الاحتمال المطلوب هو $P(-0.47 < Z < 0.95)$ وهو يساوي المساحة المظللة في

شكل (٦-٦) • ونظرا لتمثيل المنحنى الطبيعي فإن :

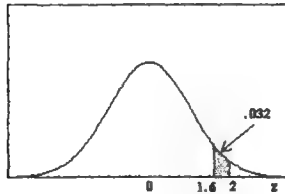
$$\begin{aligned}
 P(-0.47 < Z < 0.95) &= P(-0.47 < Z < 0) + P(0 < Z < 0.95) \\
 &= P(0 < Z < 0.47) + P(0 < Z < 0.95) = 0.1808 + 0.3289 = 0.5097.
 \end{aligned}$$



شكل (٦-٦)

(د) الاحتمال المطلوب هو $P(1.6 < Z < 2)$ وهو يساوي المساحة المظلة في شكل (٦-٧) .
أي أن :

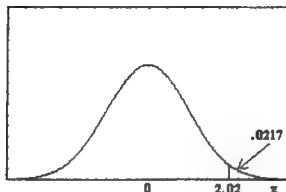
$$\begin{aligned}
 P(1.6 < Z < 2) &= P(0 < Z < 2) - P(0 < Z < 1.6) \\
 &= 0.4772 - 0.4452 = 0.032.
 \end{aligned}$$



شكل (٦-٧)

(هـ) من شكل (٦-٨) وباستخدام حقيقة أن $P(Z > 0)$ يساوي نصف المساحة الكلية تحت المنحنى الطبيعي أي أن $P(Z > 0) = 0.5$ وعلى ذلك :

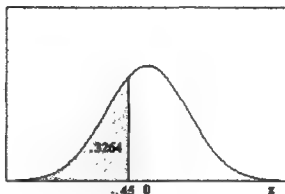
$$\begin{aligned}
 P(Z > 2.02) &= P(Z > 0) - P(0 < Z < 2.02) \\
 &= 0.5 - 0.4783 = 0.0217.
 \end{aligned}$$



شكل (٨-٦)

(و) الاحتمال المطلوب هو $P(Z < -0.45)$ وهو يساوي المساحة المظللة في شكل (٩-٦) ونظرا لمتماثل المنحنى الطبيعي فإن :

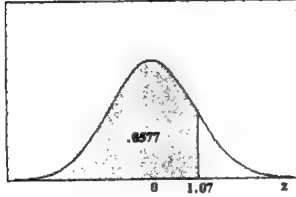
$$\begin{aligned}
 P(Z < -0.45) &= P(Z > 0.45) \\
 &= P(Z > 0) - P(0 < Z < 0.45) \\
 &= 0.5 - 0.1736 = 0.3264.
 \end{aligned}$$



شكل (٩-٦)

(ز) الاحتمال المطلوب هو $P(Z < 1.07)$ وهو يساوي المساحة المظللة في شكل (١٠-٦). ولأن المنحنى الطبيعي متماثل ومساحة كل جانب من جانبي المنحنى تساوي 0.5 فإن :

$$P(Z < 1.07) = P(Z < 0) + P(0 < Z < 1.07) \\ = 0.5 + 0.3577 = 0.8577.$$



شكل (٦-١٠)

استخدام التوزيع الطبيعي القياسي لاستخراج المساحات تحت المنحنى الطبيعي

الآن نعود مرة أخرى إلى حالة متغير عشوائي طبيعي متوسطه μ وانحرافه المعياري σ . يمكن تحويل المتغير X إلى متغير طبيعي قياسي Z باستخدام الصيغة التالية :

$$Z = \frac{X - \mu}{\sigma}$$

التحويل من X إلى Z يمثل انتقال لنقطة الأصل مصحوبا بتغير لمقياس الرسم. عندما $\mu = x$ فإن $z=0$ ، وعندما $x = \mu - \sigma$ فإن $z = -1$ ، وعندما $x = \mu + 2\sigma$ فإن $z=2$ وهكذا. أي أن مقياس الرسم قد تغير حيث تناظر مسافة σ على محور السينات مسافة قدرها واحد على محور z ، وعلى ذلك يمكن استخدام جداول التوزيع الطبيعي القياسي في ملحق (٣) لحساب الاحتمالات لأي متغير عشوائي طبيعي كما يتضح من الأمثلة التالية .

مثال (٦-١٩) في مدينة صغيرة وجد أن أعلى درجة حرارة مسجلة يوميا خلال فصل الربيع لها متوسط 20°C وانحراف معياري 5°C . بفرض أن المتغير العشوائي X (أعلى درجة حرارة يوميا) يخضع للتوزيع الطبيعي، أوجد النسبة المئوية للأيام التي فيها أعلى درجة حرارة :

(أ) بين 22°C و 26°C

(ب) على الأقل 28°C

الحل . (أ) إذا كان X يرمز لأعلى درجة حرارة مسجلة يوميا فإن X يكون متغير عشوائي طبيعي متوسطه $\mu = 20$ وانحرافه المعياري $\sigma = 5$. المتغير الطبيعي القياسي المناظر هو :

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 20}{5}$$

عندما $x_1 = 22$ فإن :

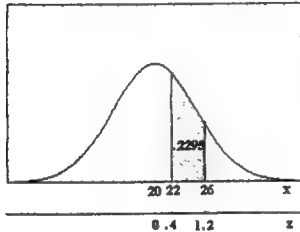
$$z_1 = \frac{22 - 20}{5} = 0.4$$

وعندما $x_2 = 26$ فإن :

$$z_2 = \frac{26 - 20}{5} = 1.2$$

الاحتمال المطلوب هو $P(22 < X < 26)$ وهو يساوي المساحة المظللة في شكل (١١-٦) .
أى أن :

$$\begin{aligned} P(22 < X < 26) &= P(0.4 < Z < 1.2) \\ &= 0.3849 - 0.1554 = 0.2295. \end{aligned}$$



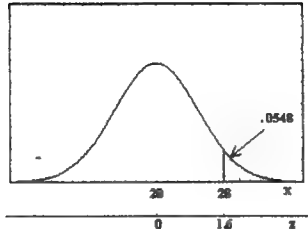
شكل (١١-٦)

أى أن النسبة المئوية للأيام التي فيها أعلى درجة حرارة بين 22°C و 26°C هي % 22.95 .
(ب) عندما $x_1 = 28$ فإن :

$$z_1 = \frac{28 - 20}{5} = 1.6$$

الاحتمال المطلوب هو $P(X \geq 28)$ وهو يساوي المساحة المظللة في شكل (١٢-٦) . أى أن

$$\begin{aligned} P(X > 28) &= P(Z > 1.6) \\ &= P(Z > 0) - P(0 < Z < 1.6) \\ &= 0.5 - 0.4452 = 0.0548. \end{aligned}$$



شكل (٦-١٢)

أى أن النسبة المئوية للأيام التي فيها أعلى درجة حرارة فوق 28°C هي 5.48 % .

مثال (٦-٢٠) إذا كانت مبيعات الغاز الطبيعي في الأسبوع في محطة لتعبئة الغاز يخضع للتوزيع الطبيعي بمتوسط $\mu = 3000$ جالون وانحراف معياري $\sigma = 200$ جالون. أوجد الاحتمال أن المبيعات في الأسبوع ما بين 3200 و 3500 جالون .
الحل . إذا كان X متغير عشوائي يرمز لمبيعات الغاز الطبيعي في الأسبوع ، فإن X يكون متغير عشوائي طبيعي متوسطه $\mu = 3000$ وانحرافه المعياري $\sigma = 200$.
عندما $x_1 = 3200$ فإن :

$$z_1 = \frac{3200 - 3000}{200} = 1.0.$$

وعندما $x_2 = 3500$ فإن :

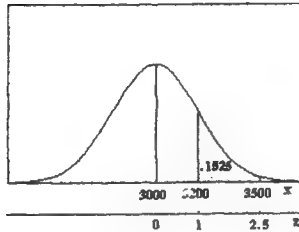
$$z_2 = \frac{3500 - 3000}{200} = 2.5.$$

الاحتمال المطلوب هو :

$$P(3200 < X < 3500)$$

وهو يساوي المساحة المظلة في شكل (٦-١٣) أى أن :

$$\begin{aligned}
 P(3200 < X < 3500) &= P(1.0 < Z < 2.5) \\
 &= P(0 < Z < 2.5) - P(0 < Z < 1) \\
 &= 0.4938 - 0.3413 = 0.1525.
 \end{aligned}$$



شكل (٦-١٣)

مثال (٦-٢١) إذا كانت كمية المطر الذي يسقط سنويا في منطقة معينة متغير عشوائي يخضع للتوزيع الطبيعي بانحراف معياري $\sigma = 2$ بوصة. أوجد المتوسط السنوي لسقوط المطر في عمام محدد إذا كان احتمال سقوط أكثر من 30 بوصة من المطر في هذا العام يساوي 0.0548 .
الحل . بفرض أن X متغير عشوائي طبيعي له متوسط غير معروف وانحراف معياري $\sigma = 2$ ، المتغير الطبيعي القياسي المناظر يكون :

$$Z = \frac{X - \mu}{\sigma}.$$

عندما $x_1 = 30$ فإن :

$$z_1 = \frac{30 - \mu}{2}.$$

ولكن احتمال أن X أكثر من 30 هو 0.0548 ، كما هو معطى، وعلى ذلك فإن :

$$P(Z \geq \frac{30 - \mu}{2}) = 0.0548.$$

أي أن :

$$P(0 \leq Z \leq \frac{30 - \mu}{2}) = 0.4452.$$

ولكن من جدول التوزيع الطبيعي القياسي المعطى في ملحق (٣) فإن :

$$P(0 \leq Z \leq 1.6) = 0.4452.$$

أي أن $\frac{30 - \mu}{2} = 1.6$ ومنه $\mu = 26.8$ بوحدة .

مثال (٦-٢٢) إذا كانت حموضة الدم الادمي مقياس بدلالة الأس الأيدروجيني مغير عشوائي طبيعي متوسطه $\mu = 7.2$. إذا كان احتمال أن يكون مستوى الأس الأيدروجيني أكبر من 7.5 يساوي 0.0222 أوجد الانحراف المعياري للتوزيع .
الحل : عندما $x_1 = 7.5$ فإن :

$$z_1 = \frac{7.5 - 7.2}{\sigma} = \frac{0.3}{\sigma}.$$

أي أن :

$$P(Z \geq \frac{0.3}{\sigma}) = 0.0222.$$

وبالتالي فإن :

$$P(0 \leq Z \leq \frac{0.3}{\sigma}) = 0.5 - 0.0222 = 0.4778.$$

ولكن من الجدول الطبيعي القياسي في ملحق (٣) نجد أن :

$$P(0 \leq Z \leq 2.01) = 0.4778.$$

أي أن $\frac{0.3}{\sigma} = 2.01$ ومنه $\sigma = 0.149$.

(٦-٦-٢) التوزيع الطبيعي كتقريب لذي الحدين

يتم الحصول على الاحتمالات المرتبطة بتجارب ذي الحدين مباشرة من الصيغة $b(x; n, p)$ أو من الجدول في ملحق (١) وذلك عندما تكون n صغيرة كما سبق ذكره في البند (٦-٢) . إذا كانت n غير موجودة في الجدول فإنه يمكن حساب احتمالات ذي الحدين بالطرق التقريبية . تناولنا في البند (٦-٤) كيفية استخدام توزيع بواسون لتقريب احتمالات ذي الحدين عندما تكون n كبيرة و p تقترب جدا من الصفر .

يعطي التوزيع الطبيعي بمتوسط $\mu = np$ وتباين $\sigma^2 = npq$ تقريبا ممتاز لتوزيع ذي الحدين عندما تكون n كبيرة و p تقترب من 0.5 . في الحقيقة يمكن استخدام التوزيع الطبيعي بمتوسط $\mu = np$ وتباين $\sigma^2 = npq$ كتقريب لتوزيع ذي الحدين حتى ولو كانت n ليست كبيرة جدا و p ليست قريبة من 0 أو 1 . وعلى أي حال، فإن هذا التقريب يكون جيّدا إذا

كانت كلا من np و nq أكبر من 5 . للتوضيح، بفرض أن X متغير عشوائي يتبع ذي الحدين حيث $n=20$ و $p=0.6$. شكل (٦-١٤) يوضح المدرج الاحتمالي لهذا المتغير العشوائي وأيضا المنحنى الطبيعي بمتوسط وانحراف معياري:

$$\sigma = \sqrt{(20)(0.6)(0.4)} = 2.19, \mu = (20)(0.6) = 12.$$

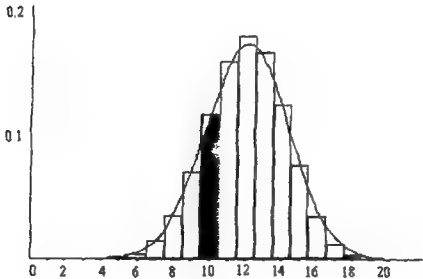
المتغير العشوائي X يتبع توزيع ذي الحدين فإن $P(X=x)$ يمثل مساحة المستطيل والذي تقع القيمة x في منتصف قاعدته وعلى ذلك يكون $b(x; n, p)$ مساو تقريبا للمساحة تحت

المنحنى الطبيعي الواقعة بين القيمتين $x - \frac{1}{2}$ و $x + \frac{1}{2}$. على سبيل المثال لحساب $P(X=10)$ وباستخدام صيغة توزيع ذي الحدين فإن :

$$\begin{aligned} P(X=10) &= \sum_{x=0}^{10} b(x; 20, 0.6) - \sum_{x=0}^9 b(x; 20, 0.6) \\ &= 0.245 - 0.128 = 0.117. \end{aligned}$$

نفس الاحتمال السابق تقريبا يساوى المساحة المظللة تحت المنحنى الطبيعي بين

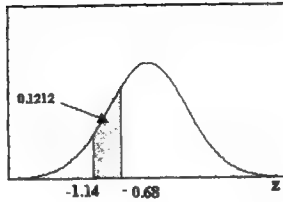
$$x_1 = 9.5 \text{ و } x_2 = 10.5 \text{ كما في شكل (٦-١٤) .}$$



شكل (٦-١٤)

أي أنه عندما $x_1 = 9.5$ فإن :

$$z_1 = \frac{9.5 - 12}{2.19} = -1.14.$$



شكل (٦-١٥)

وعندما $x_2 = 10.5$ فإن :

$$z_2 = \frac{10.5 - 12}{2.19} = -0.68.$$

إذا كان X متغير عشوائي يتبع ذي الحدين Z متغير عشوائي يتبع التوزيع الطبيعي القياسي فإن :

$$\begin{aligned} P(X = 10) &= b(10; 20, 0.6) \approx P(-1.14 \leq Z \leq -0.68) \\ &= 0.3729 - 0.2517 = 0.1212. \end{aligned}$$

والتي تتفق مع القيمة المضبوطة التي حصلنا عليها باستخدام صيغة ذي الحدين (0.117).

مثال (٦-٢٣) إذا أُلقيت زهرتا نرد 120 مرة وإذا كان اهتمامنا بمجموع النقاط التي تظهر

على السطح العلوي للتردين . فما هو احتمال ظهور المجموع 7 على الأقل 15 مرة .

الحل . بفرض أن المحاولات مستقلة وبما أن احتمال ظهور مجموع 7 في المحاولة الواحدة هو

$$\frac{6}{36} = \frac{1}{6} . \text{ فإذا كان } X \text{ متغير عشوائي يتبع ذي الحدين ويمثل عدد مرات ظهور المجموع 7 .}$$

لأن $x = 0, 1, \dots, 120$ وعلى ذلك فإن الاحتمال المطلوب هو :

$$P(X \geq 15) = \sum_{x=15}^{120} b(x; 120, \frac{1}{6}).$$

ولصعوبة حساب قيمة هذا الاحتمال فإننا نلجأ إلى التوزيع الطبيعي، كتقريب لتوزيع ذي

الحدين، والذي متوسطه وانحرافه المعياري على التوالي :

$$\mu = np = (120)\left(\frac{1}{6}\right) = 20 ,$$

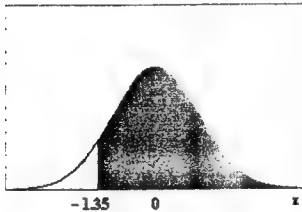
$$\sigma = \sqrt{npq} = \sqrt{(120)\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)} = \sqrt{\frac{50}{3}} = 4.08.$$

• للحصول على الاحتمال المطلوب فإننا نحتاج إلى حساب المساحة على يمين القيمة $x_1 = 14.5$ عندما $x_1 = 14.5$ فإن :

$$z_1 = \frac{14.5 - 20}{4.08} = -1.35$$

الاحتمال المطلوب يساوي المساحة المظلة في شكل (٦-١٦) ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :

$$\begin{aligned} P(X \geq 15) &= \sum_{x=15}^{120} b(x; 120, \frac{1}{6}) \approx P(Z \geq -1.35) \\ &= 0.5 + P(-1.35 \leq Z \leq 0) \\ &= 0.5 + P(0 \leq Z \leq 1.35) \\ &= 0.5 + 0.4115 = 0.9115. \end{aligned}$$



شكل (٦-١٦)

مثال (٦-٢٤) إذا كانت نسبة المصابين بارتفاع ضغط الدم في مجتمع كبير 4%. فإذا تم اختيار 500 شخصاً عشوائياً من هذا المجتمع وتم قياس ضغط دمهم، فما هو احتمال وجود 15 شخصاً على الأقل مصابين بارتفاع ضغط الدم.

الحل . إذا كان X هو عدد الأشخاص المصابين بارتفاع ضغط الدم فإنه يكون عبارة عن متغير عشوائي متقطع يتبع توزيع ذي الحدين حيث :

$$b(x; 500, 0.04) = \binom{500}{x} (0.04)^x (0.96)^{500-x}, \quad x = 0, 1, \dots, 500.$$

والاحتمال المطلوب هو :

$$b(x \geq 15) = b(15; 500, 0.04) + b(16; 500, 0.04) + \dots + b(500; 500, 0.04)$$

ومن الصعب حساب قيمة هذا الاحتمال وبالتالي فإننا نلجأ إلى التوزيع الطبيعي، كتقريب لتوزيع

ذي الحدين، والذي متوسطه وتباينه على التوالي هما :

$$\mu = np = (500)(0.04) = 20,$$

$$\sigma^2 = npq = (500)(0.04)(0.96) = 19.2.$$

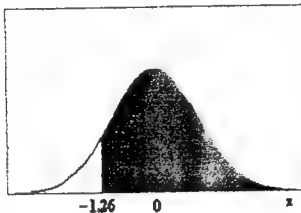
أي أن $\sigma = 4.38$ ، للحصول على الاحتمال المطلوب فإننا نحتاج إلى حساب المساحة على يمين

القيمة $x_1 = 14.5$ ، عندما $x_1 = 14.5$ فإن :

$$z_1 = \frac{14.5 - 20}{4.38} = -1.26.$$

الاحتمال المطلوب يساوي المساحة المظللة في شكل (٦-١٧) ومن جدول التوزيع الطبيعي

القياسي في ملحق (٣) فإن :



شكل (٦-١٧)

$$\begin{aligned}
 P(X \geq 15) &\approx P(Z \geq -1.26) \\
 &= 0.5 + P(-1.26 \leq Z \leq 0) \\
 &= 0.5 + P(0 \leq Z \leq 1.26) \\
 &= 0.5 + 0.3962 = 0.8962.
 \end{aligned}$$

مثال (٢٥-٦) إذا كان من المعروف أن 6% من الأفراد الذكور مصابون بعمى الألوان . فإذا تم اختيار عينة من 200 فرد من الذكور وتم اختبارهم لمعرفة إصابتهم بعمى الألوان من عدمه . أوجد احتمال أن يكون عدد المصابين بعمى الألوان :

(أ) على الأقل 20 فردا (ب) على الأكثر 15 فردا (جـ) بالضبط 15 فردا
 (د) $P(15 \leq X \leq 20)$

الحل . (أ) سوف نستخدم التوزيع الطبيعي، كتقريب لذي الحدين، بمتوسط

$$\mu = np = (200)(0.06) = 12$$

وتباين

$$\sigma^2 = npq = (200)(0.06)(0.94) = 11.28.$$

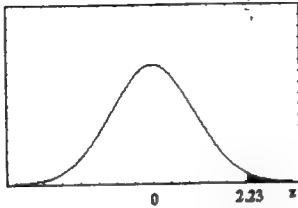
أي أن $\sigma = 3.36$. (١) للحصول على الاحتمال $P(X \geq 20)$ فإننا نحتاج إلى حساب المساحة على يمين القيمة $x_1 = 19.5$ • عندما $x_1 = 19.5$ فإن :

$$z_1 = \frac{19.5 - 12}{3.36} = 2.23.$$

الاحتمال المطلوب يساوي المساحة المظلة في شكل (٦-١٨) ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :

$$\begin{aligned}
 P(X \geq 20) &= \sum_{x=20}^{200} b(x; 200, 0.06) \\
 &\approx P(Z \geq 2.23) \\
 &= 0.5 - P(0 \leq Z \leq 2.23) \\
 &= 0.5 - 0.4871 = 0.0129.
 \end{aligned}$$

شكل (٦-١٨)

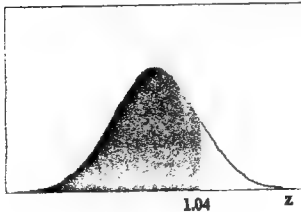


(ب) للحصول على الاحتمال $P(X \leq 15)$ فإننا نحتاج إلى حساب المساحة على يسار القيمة $x_1 = 15.5$ ، عندما $x_1 = 15.5$ فإن :

$$z_1 = \frac{15.5 - 12}{3.36} = 1.04.$$

الاحتمال المطلوب يساوي المساحة المظلمة في شكل (٦-١٩) . ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :

$$\begin{aligned} P(X \leq 15) &= \sum_{x=0}^{15} b(x; 200, 0.06) \\ &\approx P(Z \leq 1.04) \\ &= 0.5 + p(0 \leq Z \leq 1.04) \\ &= 0.5 + 0.3508 = 0.8508. \end{aligned}$$



شكل (٦-١٩)

$$P(X=15) = \sum_{x=0}^{15} b(x; 200, 0.06) - \sum_{x=0}^{14} b(x; 200, 0.06). \quad (\text{جـ})$$

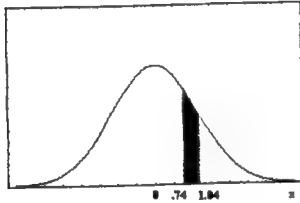
نفس الاحتمال السابق تقريبا يساوى المساحة تحت المنحنى الطبيعي بين $x_1 = 14.5$ و $x_2 = 15.5$ عندما $x_1 = 14.5$ فإن :

$$z_1 = \frac{14.5 - 12}{3.36} = 0.74.$$

وعندما $x_2 = 15.5$ فإن :

$$z_2 = \frac{15.5 - 12}{3.36} = 1.04.$$

الاحتمال المطلوب يساوى المساحة المظللة في شكل (٦-٢٠) ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :



شكل (٦-٢٠)

$$\begin{aligned} P(X=15) &\approx P(0.74 \leq Z \leq 1.04) \\ &= P(0 \leq Z \leq 1.04) - P(0 \leq Z \leq 0.74) \\ &= 0.3508 - 0.2704 = 0.0804. \end{aligned}$$

$$P(15 \leq X \leq 20) = \sum_{x=0}^{20} b(x; 200, 0.06) - \sum_{x=0}^{14} b(x; 200, 0.06). \quad (\text{د})$$

نفس الاحتمال السابق تقريبا يساوى المساحة تحت المنحنى الطبيعي بين $x_1 = 14.5$ و $x_2 = 20.5$ عندما $x_1 = 14.5$ فإن :

$$z_1 = \frac{14.5 - 12}{3.36} = 0.74.$$

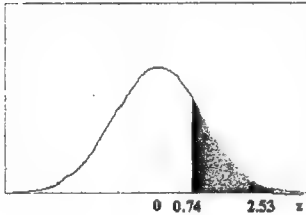
وعندما $x_2 = 20.5$ فإن :

$$z_2 = \frac{20.5 - 12}{3.36} = 2.53.$$

الاحتمال المطلوب يساوي المساحة المظلمة في شكل (٦-٢١) ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن :

$$\begin{aligned} P(15 \leq X \leq 20) &\approx P(0.74 \leq Z \leq 2.53) \\ &= P(0 \leq Z \leq 2.53) - P(0 \leq Z \leq 0.74) \\ &= 0.4943 - 0.2704 = 0.2399. \end{aligned}$$

شكل (٦-٢١)



تمارين

١- إذا ألقيت زهرة نرد متزنة مرة واحدة، وإذا كان X يمثل عدد النقاط التي تظهر على الوجه

العلوي للزهرة عند الرمي • المطلوب (أ) التوزيع الاحتمالي للمتغير العشوائي X •

(ب) $P(X > 4)$, $P(X - 3 < 0)$ (ج) التوقع والتباين للمتغير العشوائي X •

(د) التمثيل البياني للتوزيع الاحتمالي للمتغير العشوائي X •

٢- فصل دراسي يتكون من 70 تلميذاً بحيث يمثل كل تلميذ عدداً من واحد إلى 70 • بفرض

أنه تقرر اختيار تلميذ عشوائياً، وإذا كان X يمثل العدد الذي يتم إختياره • أوجد الاحتمالات الآتية

(أ) $P(X \leq 8)$ (ب) $P(5 \leq X \leq 30)$ •

٣- أوجد صيغة للتوزيع الاحتمالي الخاص بالمغير العشوائي X الذي يمثل رقم الكرة المختارة عشوائيا من وعاء يحتوي على 10 كرات مرقمة من واحد إلى 10. ما هو الاحتمال أن يكون الرقم المختار أقل من 5 ؟

٤- أوجد التوزيع المنتظم للعينات من الحجم $n=3$ المراد اختيارها من مجتمع حجمه $N=6$.

٥- أوجد التوزيع الاحتمالي للفئات الجزئية من الأشهر من الحجم $n=3$.

٦- إذا كان X متغير عشوائي منفصل يتبع التوزيع المنتظم، أي أن

$$f(x; N) = \frac{1}{N}, x = 1, 2, \dots, N$$

أوجد $E(X)$ و $\text{Var}(X)$.

٧- إذا اختر رقما عشوائيا من بين الأعداد 0, 1, ..., 9 (١) أوجد التوزيع الاحتمالي للعدد الذي يظهر (ب) التمثيل البياني لهذا التوزيع.

٨- لوحظ لفترة طويلة أن صياد يصيب الهدف باحتمال 0.9 ، فإذا أطلق الصياد 4 طلقات على هدف، أوجد احتمال (١) إصابة الهدف 3 مرات؟ (ب) إصابة الهدف مرتين على الأكثر.

٩- خلال فترة طويلة من الزمن وجد أن فاعلية دواء في شفاء الحالات التي يوصف لها هو 20% . إذا وصف هذا الدواء لأربعة مرضى . أوجد احتمال (أ) يكون له تأثير في الشفاء لثلاثة مرضى على الأقل.

١٠- أثبتت إحصائيات إحدى شركات التأمين أن 0.002 من الحوادث المسجلة في الشركة خطيرة . احسب : (١) احتمال عدم وجود حوادث خطيرة في 40 حادثة مسجلة تم اختيارها عشوائيا (ب) احتمال وقوع ثلاثة حوادث خطيرة في 20 حادثة مسجلة تم اختيارها عشوائيا .

١١- احتمال أن تصيب أي طائرة أحد أهداف العدو هو 0.9، فإذا أغارت خمس طائرات على الهدف أوجد (١) التوزيع الاحتمالي لعدد الطائرات التي تصيب الهدف (ب) متوسط التوزيع وكذلك الانحراف المعياري .

١٢- إذا كان 90% من الطلاب في جامعة ما يتنجحون في مادة الإحصاء، فما احتمال فشل اثنين في فصل يتضمن ثلاثين طالبا ؟

١٣- إذا كانت الحركات الأربعة لطائرة تعمل مستقلة بعضها عن بعض . فإن احتمال تمطيل أي منها والطائرة في الجو هو 0.5 . ما احتمال أن يكون الطيران ناجح إذا كانت عملية الطيران تحتاج تحرك واحد على الأقل ؟

١٤- إذا كان احتمال أن 4% من الوحدات في مصنع ما تالفة . ما هو احتمال إنه على الأكثر توجد وحدة تالفة في عينة عشوائية من الحجم $n=30$ ؟

١٥- احتمال أن يكسب فريق ما في أي مباراة يلعبها 0.75 فإذا كان الفريق سوف يلعب 8 مباريات أوجد احتمال أن يكسب (١) مباريتين بالضبط (ب) على الأقل مباراة واحدة (ج) أكثر من نصف المباريات .

١٦- أسرة بها خمسة أطفال . أوجد احتمال أن يكون بينهم (١) 3 أولاد (ب) عدد الأولاد أقل من عدد البنات وذلك تحت فرض أن احتمال أن يكون الطفل ولد هو 0.5 .

١٧- احتمال أن لاعب كرة القاعدة يحصل على ضربة في أي وقت يلعب فيه هو 0.3 . فإذا

كان سوف يلعب 100 مباراة في الشهر القادم أوجد المتوسط والتباين للمتغير العشوائي X .

١٨- إذا كانت فاعلية مبيد ملعن عنه في قتل الحشرات في رشه واحدة هو 90% فإذا كان

هناك 10,000 حشرة سوف تعامل برشة واحدة من هذا المبيد . أوجد التوقع الرياضي والتباين

والانحراف المعياري لعدد الحشرات التي تقتل .

١٩- قدرت شركة للطيران احتمال وصول طائرتها في ميعادها والتي تقوم من البلد A إلى البلد

B هو 0.96 . فإذا قامت خمس طائرات لهذه الشركة من مطار البلد A متجه إلى البلد B أوجد

(١) التوزيع الاحتمالي لعدد الطائرات التي تصل في ميعادها (ب) احتمال وصول طائرتين في

ميعادهما (ج) التمثيل البياني لهذا التوزيع .

٢٠- اشترى تاجر عشر تلابجات . فإذا كان احتمال أن تكون أي منها تالفة هو 0.1 فما هو

احتمال إن يكون من بينها (١) ثلاثتين تالفتين (ب) ثلاثتين على الأقل تالفتين (ج) أربعة على

الأقل تالفة .

٢١- تعمل 10 ماكينات في مصنع، فإذا كان احتمال أنه في نهاية اليوم سوف تعطل الماكينة هو

0.2 . وإذا كانت الماكينات تعمل بصورة مستقلة عن بعضها أوجد التوزيع الاحتمالي لعدد

الماكينات التي تعمل حتى نهاية اليوم مع تمثيل التوزيع الاحتمالي بالمخرج التكراري .

٢٢- بفرض أنه يوجد من بين كل 400 سيارة من إنتاج مصنع معين لإنتاج السيارات 20 سيارة

غير سليمة، أي غير صالحة للاستعمال . سحبت عينة عشوائية مكونة من 3 سيارات من إنتاج

ذلك المصنع، أوجد احتمال أن يكون من بينها سيارتين غير سليمة .

٢٣- إذا كان احتمال حصول أبوين على طفل أشقر الشعر هو 0.25، فإذا كان في الأسرة 8

أطفال . ما احتمال أن يكون نصفهم ذو شعر أشقر .

٢٤- يأخذ مراقب جودة إنتاج عينة عشوائية من 10 صمامات عشوائية من شحنة كبيرة من

الصمامات معروف عنها أنها تحتوي على 30% صماما معيما . ما هو احتمال أن يكون عدد

الصمامات المعيبة في العينة أكثر من أو يساوي 2 ؟

-٢٥- إذا كان احتمال إصابة هدف بقليلة واحدة هو 0.2، ما هو احتمال إصابة الهدف مرتين على الأقل خلال 3 قذائف ؟

-٢٦- احتمال كشف جهاز رادار لطائرة معادية هو 0.99، فإذا كان لدينا 4 أجهزة تعمل مستقلة بعضها عن بعضها، المطلوب إيجاد (أ) ظهور طائرة معادية في سماتنا (ب) ظهور طائرة معادية على شاشات أربعة منها .

-٢٧- إذا كان 5% من نوع معين من صمامات التلفزيون يحترق قبل انتهاء مدة الكفالة . فإذا بيع ألف صمام فما هو متوسط وتباين \bar{X} ، حيث \bar{X} تمثل عدد الصمامات المحترقة قبل مدة كفالتها ؟

-٢٨- إذا كان احتمال تحمل نوع معين من المصاييح للضغط العالي هو 0.1 فإذا أخذنا عينة من 200 مصباح فما هو احتمال ألا يتحمل 20 منها للضغط العالي؟

-٢٩- في مدينة كبيرة كان عدم الوفاق بين الزوجين سبباً لـ 60% من حالات الطلاق . أوجد احتمال أن أربعة من حالات الطلاق الستة القادمة في المدينة سيعزى إلى هذا السبب .

-٣٠- في بحث ميداني في بلد ما وجد أن 2% من الأشخاص يفضلون شراء تلفزيون بلون أبيض . ما هو احتمال أن أكثر من نصف التلفزيونات التي سوف تباع من بين 40 تلفزيون لونها أبيض ؟

-٣١- يحتوي امتحان على 18 سؤال من نوع صح وخطأ، كل سؤال له أربع أجوبة محتملة منهم فقط واحد هو الإجابة الصحيح . ما هو احتمال أن شخص يمكنه الإجابة على 5 إجابات صحيحة وذلك بطريقة التخمين ؟

-٣٢- احتمال أن يشفى مريض من عملية في القلب هو 0.9 ، ما هو احتمال أن يشفى 5 مرضى من بين 10 مرضى سوف يجري لهم العملية ؟

-٣٣- يحتوي امتحان على 20 سؤال من نوع صح وخطأ، كل سؤال له أربع أجوبة محتملة منهم فقط واحد هو الإجابة الصحيحة . فإذا كان احتمال أن يحصل طالب على 100% في الاختبار

$$\text{هو } P(100\%) = \left(\frac{1}{4}\right)^{20} \text{ . المطلوب إيجاد احتمال حصوله على } 80\% \text{ إجابة صحيحة .}$$

-٣٤- إذا كان 0.77 من المستهلكين يستخدمون مسحوق ما . ما هو الاحتمال في عينة من 10 مستهلكين (أ) بالضبط أربعة يستخدمون المسحوق (ب) على الأقل أربعة يستخدمون المسحوق (ج) ليس أكثر من واحد يستخدم المسحوق (د) 8 أو أكثر يستخدمون المسحوق .

-٣٥- أثبتت الإحصائيات في إحدى المستشفيات أن 40% من المرضى الذين يتلقون العلاج بها يفشلون في دفع فاتورة الحساب . بفرض أن 4 مرضى دخلوا المستشفى لتلقى العلاج، أوجد

الاحتمالات التالية (أ) كل المرضى لن يدفعون الفاتورة (ب) واحد على الأقل سوف يدفع الفاتورة .

٣٦- إذا كان 90% من المتقدمين لوظيفة دبلوماسي يفشلون في المقابلة الأولى، فإذا اختبر 12 شخصا عشوائيا من المتقدمين، أوجد احتمال انه (أ) على الأقل 5 يفشلون (ب) من 10 إلى 12 يفشلون .

٣٧- إذا كان X متغيرا عشوائيا يمثل عدد الطلبة الذين يحصلون على وظائف بعد سنة واحدة من التخرج يتبع توزيع ذي الحدين بمعلمتين $n=250$, $p=0.98$ (أ) أوجد التوزيع الاحتمالي للمتغير العشوائي X (ب) أوجد المتوسط والتباين للتوزيع الاحتمالي للمتغير X .

٣٨- أثبتت التجارب حدوث تلف في التمثيل الغذائي لطفل واحد من بسين 100 يولدون . بفرض أنه تم ولادة أربعة أطفال في يوم معين، ما هو احتمال (أ) ليس أكثر من طفل واحد لديه تلف (ب) عدم حدوث تلف .

٣٩- بفرض أن 20% من سائقي سيارات الأجرة الذين يصلون إلى موقف السيارات يعطون إشارات حمراء في كل الاتجاهات (يتبعون الأنظمة) فإذا اختبر 20 سائقا عشوائيا من الذين تم وصولهم إلى موقف السيارات . أوجد احتمال (أ) على الأكثر 6 يتبعوا النظام (ب) بالضبط 10 يتبعون النظام (ج) على الأقل 12 يتبعون النظام .

٤٠- استخدم جدول ذي الحدين في جدول ذي الحدين في ملحق (أ) في إيجاد الاحتمالات

$$P(x = 4), n = 15, p = 0.4$$

$$P(x \leq 4), n = 20, p = 0.1$$

$$P(5 \leq X \leq 11), n = 20, p = 0.2$$

$$P(x = 6), n = 15, p = 0.5$$

$$P(x \geq 8), n = 15, p = 0.5$$

٤١- في إحدى الدراسات لولاء مؤسسات الأحداث وجد أن 90% من الولاء قد يعودون مرة ثانية إلى المؤسسة بعد انتهاء مدقم . فإذا أخذت عينة عشوائية مكونة 5 من نزلاء من إحدى المؤسسات ، أوجد (أ) التوزيع الاحتمالي لعدد الولاء الذين يعودون مرة أخرى إلى المؤسسة . (ب) عدد الولاء المتوقع عودهم (ب) الانحراف المعياري لعدد الولاء المتوقع عودهم مرة أخرى إلى المؤسسة .

٤٢- إذا كان X متغيرا عشوائيا يمثل عدد المرضى الذين يلغون حجزهم عند طبيب الأسنان، فإذا كان معلم توزيع ذي الحدين للمتغير X هو $n=15$, $p=0.05$ ، أوجد احتمال (أ) ثلاثة أشخاص سوف يلغون الحجز (ب) على الأكثر أربعة سوف يلغون الحجز .

-٤٣- في تجربة زراعية كانت نسبة الإصابة بفطر ما 0.2 في نهاية التجربة . فإذا كان X مصفراً عشوائياً يمثل عدد النباتات المصابة بالفطر في عينة من 10 نباتات، أوجد (أ) التوزيع الاحتمالي لعدد النباتات المصابة (ب) أوجد $P(X=0)$.

-٤٤- يولد 30% من مواليد سلالة معينة من الأرناب بشعر طويل أوجد (أ) التوزيع الاحتمالي لعدد الحيوانات التي تولد بشعر طويل في بطن من أربعة أرناب (ب) أوجد $P(X=1)$ (ج) أوجد المتوسط والتباين للتوزيع الاحتمالي للمتغير X .

-٤٥- يبت 75% من شاتلات الأشجار المزروعة بشركة معينة لتشجير الطرق أوجد المتوسط والتباين للتوزيع الاحتمالي للمتغير X والذي يمثل عدد الأشجار التي تبت من مجموعة مكونة من 10 أشجار مزروعة .

-٤٦- أعطيت ستة فئران جرعة معينة من سم ولوحظ عدد الفئران التي قوت خلال 72 ساعة . فإذا كان احتمال موت كل فأر هو 0.2 . أوجد التوزيع الاحتمالي لعدد الفئران التي تموت خلال 72 ساعة وأوجد المتوسط والتباين للمتغير العشوائي X .

-٤٧- إذا كان لـ 33% من القرى هي 33% ، فإذا أخذت عينة عشوائية مكونة من 5 أشخاص من أفراد هذه القرية . أوجد التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل عدد الأميين في العينة وأوجد المتوسط والتباين للمتغير العشوائي X .

-٤٨- درب حيوان على لمس واحدة من رافعتين إذا أمر بذلك . بفرض أن احتمال أن يلمس الرافعة الصحيحة إذا أمر بذلك هو 0.8 أوجد التوزيع الاحتمالي لعدد مرات لمس الرافعة الصحيحة في محاولات عددها 10 وأوجد المتوسط والتباين للتوزيع الاحتمالي الذي حصلت عليه .

-٤٩- وجد أن فصيلة دم 40% من سكان مدينة ما تكون من نوع معين . ما احتمال أن يكون ثلاثة فقط من مجموعة مكونة من ثمانية أشخاص تم اختيارهم عشوائياً من سكان هذه المدينة هم فصيلة الدم هذه؟

-٥٠- بالرجوع إلى التاريخ العائلي لزوجين وجد أن احتمال أن يكون أياً من أطفالهم مصاباً بعبس خلقي معين 0.05 . فإذا كان للزوجين أربعة أطفال أوجد التوزيع الاحتمالي لعدد الأطفال المصابين بتخلف عقلي وأوجد المتوسط والتباين للمتغير العشوائي X .

-٥١- إذا كان احتمال ولادة طفل أعسر في بلد ما، هو 0.01 أوجد التوزيع الاحتمالي لعدد الأطفال الذين لديهم هذه الصفة . وأوجد المتوسط والتباين للمتغير العشوائي X .

-٥٢- إذا كان احتمال التحاق إحدى الخريجين بالعمل فور تخرجه هو 0.75 . فإذا اختيرت عينة عشوائية من 30 فرد من حديثي التخرج . أوجد (أ) احتمال عدم التحاق أي شخص بالعمل فور

تخرجه (ب) احتمال التحاق اثنين على الأقل بالعمل فور تخرجهم (جـ) العدد المتوقع للأشخاص الذين يلتحقون بالعمل فور تخرجهم .

-٥٣- في عملية تصنيع كرات التجميل وجد أن احتمال وجود كرة تالفة 0.1 ما هو احتمال الحصول على 10 كرات تحميل تالفة من عينة عشوائية لها 1000 وحدة؟

-٥٤- إذا كان طرد بريدي يمكن أن يفقد أو يتلف أو يصل إلى صاحبه . فإذا كان احتمال أن يتلف 0.2 واحتمال أن يفقد هو 0.4 واحتمال أن يصل إلى صاحبه هو 0.4 . فإذا أرسل 15 طردا إلى بلد ما هو احتمال أن يصل 13 منهم بأمان إلى أصحابهم ويفقد 1 ويتلف 1 .

-٥٥- تما لنظرية الوراثة لأن نوع معين من الحيوانات تلد حيوانات لونها أحمر وأبيض وأسود بنسبة 4 : 4 : 8 أوجد احتمال أنه من بين 8 مواليد سوف يكون منهم 5 لونهم أحمر و 2 لونهم أسود وواحد لونه أبيض .

-٥٦- صندوق به 15 ثمرة لأكهة، منها 4 تالفة، أوجد التوزيع الاحتمالي لعدد الثمار التالفة ثم مثل بيانيا الدالة الإحصائية و $3 < 8$

-٥٧- يوجد في مكتبة 20 نسخة من كتاب في مقدمة الإحصاء، منهم 12 طبعة أولى و 8 طبعة ثانية . فإذا اخبرت عينة عشوائية من الحجم $n=5$ ، وإذا كانت X تمثل عدد الكتب المختارة من الطبعة الثانية، أوجد $P(X=2)$.

-٥٨- صندوق يتلقى على رفاقتي بالغة الصغر منها 10 جيدة و 3 تالفة . فإذا تقرر اختيار عينة عشوائية من ثلاثة رفاقتي . وإذا كانت X تمثل عدد الوحدات التالفة في العينة المختارة، أوجد $P(X \leq 3)$.

-٥٩- قام باحث في علم الجيولوجيا بتجميع 10 وحدات من صخور البازلت و 10 وحدات من الجرانيت، فإذا كان للباحث معمل وطلب من مساعده أن يختار عينة عشوائية من 15 وحدة لتحليل أوجد (أ) التوزيع الاحتمالي لعدد وحدات البازلت في العينة المختارة (ب) احتمال أن الوحدات المختارة من نفس النوع .

-٦٠- إذا كان عدد الجزيئات المشعة من مصدر إشعاعي يتبع توزيع بواسون وإذا كان $P(X=0)=0.3$ فما احتمال إشعاع جزيئين أو أكثر ؟

-٦١- وعاء به جزيئات 10000 فإذا كان احتمال هروب جزيء من الوعاء هو 0.0004 فما احتمال هروب أكثر من 5 جزيئات .

-٦٢- تمر في المتوسط 20 سيارة في الدقيقة من أمام كشك رسوم المرور خلال ساعة السزروة . أوجد احتمال مرور 7 سيارات من أمام الكشك خلال دقيقة مختارة عشوائية .

-٦٣- إذا كانت نسبة المعيب في مصنع مصابيح كهربائية هو 0.1 ، فإذا اخترنا عينة عشوائية من 200 مصباح، أوجد احتمال أن يوجد في هذه العينة مصباحا واحدا على الأكثر معيب وذلك باستخدام توزيع بواسون كتقريب لذى الحدين .

-٦٤- إذا كان متوسط وقوع الزلازل في بلد ما هو 3 في السنة ، أوجد احتمال أن يقع زلزالا واحدا على الأقل في هذه السنة .

-٦٥- إذا كانت متوسط عدد الحوادث الشهرية في إحدى الطرق هو 0.2 . اختر شهرا عشوائيا، أوجد احتمال وقوع حادثتين على الأقل .

-٦٦- تقوم إحدى المصانع بإنتاج منتج معين معا في أكياس عبوة الكيس الواحد كيلوجرام . فإذا كان احتمال وجود كيس واحد غير مطابق للمواصفات هو 0.05 ، فإذا اختر عينة عشوائية من 20 كيس وبالفرض أن X متغيرا عشوائيا يمثل عدد الأكياس الغير مطابقة في العينة . أوجد (أ) التوزيع الاحتمالي لعدد الأكياس الغير مطابقة للمواصفات في العينة (ب) احتمال وجود كيسين على الأقل غير مطابقين للمواصفات (ج) احسب العدد المتوقع للأكياس الغير مطابقة للمواصفات .

-٦٧- في المتوسط يتوقف 7 عملاء للترود بالوقود عند طلمبة بولين في الساعة . ما هو احتمال (أ) توقف 7 عملاء في الساعة (ب) أربعة عملاء أو أقل في ساعة ما .

-٦٨- تشير الدراسات على أن 0.002 من القوى العاملة القومية في بلد ما يصابون بمرض خطير خلال عام . فإذا اختر شخص عشوائيا . أوجد (أ) القيمة المتوقعة لعدد الذين يمرضون في العام (ب) احتمال أن عاملين يمرضون خلال عام . استخدم توزيع بواسون كتقريب ذي الحدين .

-٦٩- بفرض أن متغيرا عشوائيا يمثل عدد الجرائم التي تحدث في بلد ما في الفترة ما بين الساعة الواحدة صباحا حتى الساعة الثانية يتبع توزيع بواسون حيث $\mu = 0.2$ (أ) أوجد التوزيع الاحتمالي للمتغير العشوائي X (ب) أوجد $P(X=0)$.

-٧٠- إذا كان عدد التذاكر التي تصرف في موقف للسيارات في الساعة يتبع توزيع بواسون بمعلمة $\mu = 4$ (أ) ما هو الاحتمال أنه بالضبط تصرف 3 تذاكر خلال ساعة معينة (ب) ما هو الاحتمال أنه على الأقل تصرف 4 خلال ساعة معينة (ج) ما هو عدد التذاكر المتوقع صرفها خلال 45 دقيقة .

-٧١- تصل طائرة شراعية إلى المطار بمعدل $\mu = 8$ في الساعة، ما هو الاحتمال (أ) خمسة بالضبط سوف يصلون خلال ساعة (ب) على الأقل خمسة يصلون خلال ساعة (ج) على الأقل 10 يصلون خلال ساعة (د) ما هي القيمة المتوقعة والانحراف المعياري لعدد الطائرات التي تصل إلى المطار خلال 90 دقيقة (هـ) ما هو الاحتمال أنه على الأقل تصل طائرة خلال 2.5 ساعة ؟

-٧٢- في بحث وجد أن شخص واحد من 1000 شخص يحمل جين تالف يسؤدى إلى الإصابة بسرطان القولون . اختبرت عينة عشوائية من 15 شخص أوجد (أ) التوزيع التقريبي لعدد الأشخاص الذين يحملون الجين التالف (ب) استخدم هذا التوزيع في حساب الاحتمال التقريبي أن 7 يحملون الجين التالف .

-٧٣- قامت شركة لبيع السيارات بإرسال طلب لكل مشر منها سيارة معينة وذلك لفحصها من العيوب . يفرض أن 0.001 يمثل احتمال وجود عيب في السيارة . إذا اختبرت عينة عشوائية 10000 أوجد (أ) القيمة المتوقعة لعدد السيارات في العينة التي بها هذا العيب (ب) ما هو الاحتمال التقريبي أنه على الأقل 10 سيارات بها هذا العيب (جـ) ما هو الاحتمال التقريبي لعدم وجود سيارات في العينة بها عيوب .

-٧٤- يستقبل متروال في المتوسط ثلاث مكالمات في الدقيقة وذلك من أماكن بعيدة والمطلوب (أ) احتمال عدم وصول أي مكالمات في دقيقة واحدة (ب) احتمال وصول مكالمتين في دقيقتين (جـ) احتمال وصول أربع مكالمات في ثلاث دقائق .

-٧٥- إذا كان X متغير عشوائي يمثل عدد الأشخاص الذين يصلون إلى نادي رياضي بمتوسط $\mu = 2$ في الدقيقة المطلوب (أ) احتمال وصول شخص واحد في دقيقة واحدة (ب) احتمال وصول 5 أشخاص في ثلاثة دقائق (جـ) احتمال عدم وصول أي شخص في دقيقة واحدة .

-٧٦- يستقبل متروال في المتوسط ثلاثة مكالمات في الدقيقة وذلك من أماكن بعيدة والمطلوب (أ) احتمال عدم وصول أي مكالمات في دقيقة واحدة (ب) احتمال وصول مكالمتين في دقيقتين (جـ) احتمال وصول 4 مكالمات في ثلاث دقائق .

-٧٧- تقوم ماكينة بتصنيع الأقمشة، فإذا كان في المتوسط يوجد عيب لكل 10 ياردة من القماش أوجد (أ) احتمال عدم وجود عيوب في 2 ياردة من القماش (ب) احتمال وجود عيب في ياردة من القماش .

-٧٨- إذا كان عدد المكالمات التي يتلقاها مركز للشرطة بين 8:30 و 8:00 مساءً في يوم الجمعة يتبع توزيع بواسون بمعلمة $\mu = 2$ أوجد (أ) احتمال عدم وصول مكالمتين خلال هذه الفترة (ب) احتمال وصول مكالمتين خلال هذه الفترة (جـ) احتمال وصول ثلاث مكالمات أو أقل خلال هذه الفترة .

-٧٩- قررت عائلة تنظيم الإنجاب إذا رزقها الله بخمسة ذكور . فإذا كان احتمال ولادة ذكر في هذه العائلة هو 0.4 . أوجد التوزيع الاحتمالي لعدد مرات الحمل (الوضع) .

٨٠- قررت أسرة أن تنظم النسل إذا رزقها الله بثلاثة أطفال من نفس الجنس . بفرض أن احتمال ولادة طفل ذكر تساوى احتمال ولادة طفل أنثى تساوى 0.5 . أوجد التوزيع الاحتمالي لعدد مرات الحمل .

٨١- يلعب فريق A مع الفريق B سلسلة من المباريات . فإذا كان احتمال أن A يكسب في المباراة الواحدة التي يلعبها 0.6 وإذا كانت المباريات مستقلة، أوجد احتمال أن الفريق A قد يكون كسب أربعة مباريات في المحاولة السادسة .

٨٢- أوجد احتمال أن شخص يلقى عملة سوف يحصل على صورة في المحاولة السابعة.

٨٣- احتمال أن طالب يجتاز امتحان للحصول على رخصة قيادة طائرة هو 0.7 أوجد احتمال أن الشخص ينجح في الامتحان (أ) في المحاولة الثالثة (ب) قبل المحاولة الخامسة .

٨٤- إذا كانت الدرجات إلى حصل عليها 100 طالب تتبع التوزيع الطبيعي بمتوسط 60 درجة وتباين 0.5. اختبر طالبا بطريقة عشوائية (أ) ما هو احتمال أن تزيد درجته عن 72 ؟ (ب) ما هو عدد الطلاب الذين تقل درجاتهم عن 56 ؟

٨٥- إذا حصل 10% من الطلاب على جوائز بسبب ارتفاع درجاتهم فما هي أدنى درجة يجب أن يحصل عليها الطالب حتى يحصل على جائزة ؟

٨٦- إذا كان التوزيع التكراري لضغط الدم طبيعيا وكان متوسط الضغط الطبيعي هو 120 سم من الزئبق والانحراف المعياري هو 15 سم من الزئبق .

(أ) فما هو نسبة الأشخاص المحتمل أن يكون ضغطهم 150 فأكثر ؟

(ب) إذا علمت أن احتمال الحصول على شخص ضغط دمه أقل من قيمة معينة هي x_1 مثلا هو 0.8942 ، فما هي قيمة x_1 ؟

٨٧- إذا كان الزمن اللازم لحضم وحلة واحدة من طعام معين يتبع التوزيع الطبيعي بمتوسط قدره 25 دقيقة وانحراف معياري قدره 3 دقائق . ما هو احتمال أن تحضم وحدة طعام في أقل من 30 دقيقة ؟

٨٨- أوزان الدجاج في مزرعة ما تتبع التوزيع الطبيعي بمتوسط 6 رطل وانحراف معياري 1.5 أوجد (أ) احتمال أن يكون وزن الدجاجة اختبرت عشوائيا أكبر من سبعة أرطال .

(ب) نسبة الدجاج الذي وزنه أقل من أربعة أرطال .

٨٩- تنتج ماكينة للمشروبات الباردة في المتوسط 7 أوقيات من العصير لكل كوب . بفرض أن كمية الشراب تتبع التوزيع الطبيعي بانحراف معياري 0.5 أوقية . المطلوب (أ) نسبة الأكواب التي تحتوي على الأقل 7.8 أوقية . (ب) ما هو احتمال أن كوب يحتوي من بين 6.7 إلى 7.9 أوقية .

٩٠- يقطع شخص المسافة من موله إلى عمله يوميا في زمن قدره 24 دقيقة في المتوسط بانحراف معياري 3.8 دقيقة . بفرض أن الزمن الذي يستغرقه يوميا يتبع التوزيع الطبيعي . أوجد احتمال أن الزمن الذي يستغرقه على الأقل نصف ساعة .

٩١- إذا كانت أطوال 1000 طالبا تتبع التوزيع الطبيعي بمتوسط 68.5 بوصة وانحراف معياري 2.7 بوصة . كم عدد الطلبة المتوقع أن تكون أطوالهم : (أ) أقل من 63 بوصة . (ب) بين 67.5 إلى 71 بوصة . (جـ) تساوي 69 بوصة . (د) أكبر من 74 بوصة .

٩٢- تدفع شركة أجور العاملين فيها بمتوسط 100 جنيه لكل ساعة بانحراف معياري 5 جنيه . إذا كانت الأجور تقريبا تتبع التوزيع الطبيعي . ما هي النسبة المئوية من للعاملين الذين أجورهم تتراوح بين 80 إلى 90 في الساعة ؟

٩٣- إذا كان معدل الذكاء $I.Q$ مجموعة من الطلبة الراغبين في الالتحاق في جامعة مما يتبع التوزيع الطبيعي بمتوسط $\mu=115$ وانحراف معياري $\sigma=12$. أوجد احتمال أن يكون معدل الذكاء أكبر من 120 .

٩٤- إذا كان متوسط العمر لمولد كهربائي صغير هو 10 سنوات بانحراف معياري 2 سنة . أوجد احتمال أن يقل عمر المولد عن 8 سنوات .

٩٥- إذا كانت درجة الحرارة في مدينة ما يتبع التوزيع الطبيعي بمتوسط 20 درجة وانحراف معياري 3 درجات . أوجد احتمال أن تكون درجة الحرارة في أحد الأيام أقل من 25 درجة .

٩٦- إذا كان قطر السلك الكهربى من إنتاج شركة ما يتبع التوزيع الطبيعي بمتوسط 8 ملليمتر وتباين 0.0004 ملليمتر . اشترى شخص سلك فما هو احتمال أن قطره لا يعدي 8.0 ملليمتر .

٩٧- إذا كان متوسط أطوال مجموعة من الجنود في معركة هو 70 بوصة . وإذا كان 10% من الجنود أطول من 72 بوصة . بفرض أن أطوال الجنود في هذه المعركة يتبع التوزيع الطبيعي ما هو الانحراف المعياري ؟

٩٨- إذا كان أطوال مجموعة من الجنود يتبع التوزيع الطبيعي . إذا كان 13.75 % منهم أطول من 72.2 بوصة و 8.08 % أقصر من 67.2 بوصة . ما هو المتوسط والانحراف المعياري لأطوال الجنود ؟

٩٩- إذا كان أوزان الذكور البالغين يتبع التوزيع الطبيعي وأن 6.6 % تحت 130 رطل وأن 77.45 % بين 130 إلى 180 رطل . أوجد معالم التوزيع .

١٠٠- إذا كان دخول الأسر في مدينة ما يتبع التوزيع الطبيعي بمتوسط $\mu=15000$ وانحراف معياري $\sigma=3000$. ما هو احتمال أن يكون دخل أسرة اختيرت عشوائيا بين :
(أ) من 16000\$ إلى 18000\$ (ب) أقل من 12000\$ (جـ) اعلى من 15000\$.

١٠١- تصب درجات اختبارات الذكاء لمطوعي الجيش في سنة ما التوزيع الطبيعي بمتوسط $\mu=110$ وانحراف معياري $\sigma=10$. ويريد الجيش أن يعطي تدريباً متقدماً لأعلى 10% من درجات اختبارات الذكاء. ما هي أقل درجة في اختبارات الذكاء التي تقبل لحضور التدريب المتقدم ؟

١٠٢- إذا أُلقيت زهرة نرد 400 مرة. استخدم التوزيع الطبيعي كتقريب لذي الحدين لإيجاد احتمال ظهور الرقم واحد (أ) 185 إلى 210 مرة (ب) بالضبط 205 مرة .

(ج) أقل من 176 مرة .

١٠٣- في مدينة ما وجد أن 10% من المدخنين مصابون بالسرطان. أخذت عينة من هذه المدينة من 300 مدخن وفحصوا للتحقق من إصابتهم بالسرطان المطلوب:

(أ) احتمال أن تحوى العينة على 25 شخص مصاب بالسرطان .

(ب) ستون شخصاً على الأقل مصاباً بالسرطان .

(ب) باستخدام التوزيع الطبيعي كتقريب لذي الحدين)

١٠٤- أُلقيت قطعة نقود 20 مرة. احسب احتمال الحصول على 8 صور باستخدام:

(أ) توزيع ذي الحدين (ب) التوزيع الطبيعي كتقريب لذي الحدين .

١٠٥- إذا كان 70 % من الطلاب المتدربين بالكليات يحصلون على مؤهلاتهم. أوجد احتمال أنه من بين 20 طالباً مختارين عشوائياً من المتدربين حديثاً سوف يحصل على أكثر من 10 طلاب منهم على المؤهل (باستخدام التوزيع الطبيعي كتقريب لذي الحدين) .

١٠٦- بفرض أن 45% من المكالمات التي يستقبلها عامل تليفون في شركة ما من مسافات بعيدة . ما هو احتمال الحصول على 11 مكالمات من مسافة بعيدة من بين 20 مكالمات يستقبلهم استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين .

١٠٧- تستهلك شركة ما كمية من الطلح يوميا في المتوسط 600 رطل بالبحراف معياري 25 رطل . فإذا كانت الكمية المستهلكة يوميا تتبع التوزيع الطبيعي . أوجد (أ) احتمال أن تستهلك الشركة أكثر من 700 رطل يوميا (ب) احتمال أن تستهلك من 500 إلى 800 يوميا .

١٠٨- إذا كانت $p=0.3$ استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين لتقدير احتمال الحصول على 140 وحدة تالفة من بين 500 وحدة تالفة منتجة من إحدى المصانع .

١٠٩- إذا كان معروف أن نقطة الذوبان للذهب هي 1.06°C (في المتوسط) بالبحراف معيولي 0.3°C أوجد احتمال $P(X>1.77)$.

١١٠- إذا كان الطلب على اللحوم في مخزن لبيع اللحوم، خلال أسبوع ، تقريبا يتبع التوزيع الطبيعي بمتوسط 5000 رطل والانحراف معياري 300 رطل أوجد $P(X>5300)$ في اسبوع .

- ١١١- إذا كان احتمال أن تنجب سيدة مولود ذكر 0.5 . فإذا كان هناك 100 سيدة في هذه القرية حوامل وبالفرض أن X ترمز لعدد المواليد الذكور أوجد (أ) التوزيع الاحتمالي للمتغير العشوائي X (ب) احتمال أن تنجب على الأقل 3 مواليد ذكور (ج) احتمال عدم إنجاب طفل ذكر (استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين) .
- ١١٢- صممت سيارة جديدة تحت فرض أن 70% من المبيعات سوف تكون للسيدات . إذا اختبرت عينة عشوائية من 500 مشرئ . ما هو احتمال أنه على الأقل 270 منهم سيدات . (استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين) .
- ١١٣- إذا كان احتمال بقاء مريض لأكثر من ثمانية وأربعين ساعة في قسم العناية المركزة بمستشفى هو 0.055 ، المطلوب إيجاد احتمال أن يمكث عشرة مرضى أكثر من ثمانية وأربعين ساعة من بين خمسين مريضاً الخفوا بالقسم في يوم محدد (استخدم التوزيع الطبيعي كتقريب لتوزيع ذي الحدين) .

الفصل السابع

توزيعات المعاينة

Sampling Distributions

(١-٧) مقدمة Introduction

يهتم فرع الإحصاء الاستدلالي بالتعميم والتنبؤ، فعلى سبيل المثال يمكن القول أن متوسط دخل الفرد في بلد ما \$ 86000 في السنة وذلك بناء على عينة عشوائية اختبرت من هذا البلد. وبتوضيح آخر يمكن أن نتوقع بناء على آراء مجموعة من الأشخاص في الشارع أن 80% من أصوات الناخبين في مدينة ما سوف تعطى لمرشح معين. كما يمكن التوقع أن عمر الصباح الكهربائي من إنتاج مصنع ما يتراوح بين 1150 ساعة و 1250 ساعة بدرجة ثقة معينة. فإننا نجد في كل مثال من الأمثلة السابقة، تم حساب إحصاء من عينة عشوائية تم اختيارها من المجتمع موضع الدراسة، ومن تلك الإحصاءات أمكننا الوصول إلى جمل تحس قيم العالم والتي قد تكون صحيحة أو غير صحيحة. التعميم من الإحصاء إلى المعلمة يكون بدقة فقط إذا استطعنا أن نفهم العلاقة بين المجتمع وعيناته.

في الفصل الخامس عرفنا الإحصاء على أنه متغير عشوائي يعتمد قيمته فقط على العينة، وبالتالي فإن نفس الحسابات لعينات مختلفة من المجتمع تؤدي إلى قيم مختلفة للإحصاء. هذه الاختلافات في قيم الإحصاء تعتمد على حجم المجتمع وحجم العينات والطريقة التي استخدمت في اختيار العينات العشوائية. إذا كان حجم المجتمع كبيراً أو لانهائي فإن التوزيع الاحتمالي للإحصاء في حالة السحب بإرجاع سوف يكون نفسه في حالة السحب بدون إرجاع. ومن ناحية أخرى فإن السحب بإرجاع من مجتمع صغير محدود يعطى توزيعاً للإحصاء يختلف قليلاً عن السحب بدون إرجاع. أخيراً المعالجة مع الإرجاع من مجتمع محدود يكافئ المعالجة من مجتمع لانهائي وذلك لعدم وجود حدود لحجم العينة المختارة من المجتمع.

تعريف: التوزيع الاحتمالي لأي إحصاء يسمى التوزيع العيني **sampling distribution**

تعريف: الانحراف المعياري للتوزيع العيني لأي إحصاء يسمى الخطأ المعياري **standard error** للإحصاء.

فعلى سبيل المثال التوزيع الاحتمالي للإحصاء \bar{X} يسمى التوزيع العيني للمتوسط، كما أن الخطأ المعياري للمتوسط هو الانحراف المعياري للتوزيع العيني للإحصاء \bar{X} .

في هذا الفصل سوف ندرس بعض توزيعات المعالجة الأكثر استخداماً في الإحصاء. التطبيقات على تلك التوزيعات العينية تخص مشاكل الاستدلال الإحصائي التي سوف نتناولها في الفصل الثامن والتاسع.

Normal Sampling Distributions

(٧-٢) توزيعات المعاينة الطبيعية

إذا أخذنا عينات متكررة من الحجم n من توزيع متصل له متوسط μ وتباين σ^2 . لكل عينة ثم حساب القيمة y لإحصاء ما Y ، والذي نفسه متغير عشوائي متصل. بفرض أن Y يتبع توزيعاً طبيعياً بمتوسط μ_Y وانحراف معياري σ_Y . النظرية التالية تنص على أن :
نظرية (٧-١) إذا كان y قيمة للإحصاء Y والذي يتبع توزيعاً طبيعياً بمتوسط μ_Y وانحراف معياري σ_Y ، فإن :

$$z = \frac{y - \mu_Y}{\sigma_Y}$$

هي قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي حيث :

$$Z = \frac{Y - \mu_Y}{\sigma_Y}$$

يمكن تطبيق نظرية (٧-١) للإحصاءات المحسوبة من عينات عشوائية اختيرت من مجتمعات متقطعة، سواء محدودة أو غير محدودة، والتي التوزيعات المعينة لإحصاءاتها تقريباً تتبع توزيعاً طبيعياً. هذا يمكننا استخدام جدول التوزيع الطبيعي في الملحق (٣) في حساب الاحتمالات التي يأخذها الإحصاء في فترات معينة.

Sampling Distributions of the Mean (٧-٣) توزيعات المعاينة للمتوسط

يعتبر توزيع المعاينة للمتوسط \bar{X} أهم توزيع معاينة سوف نتناوله في هذا الفصل. إن شكل ونوع التوزيع الاحتمالي لمجتمع متوسط العينات (التوزيع العيني للمتوسط) يعتمد على شكل المجتمع الأصلي الذي اختيرت منه العينات. النظرية الآتية تعطي التوزيع العيني للمتوسط إذا كان المجتمع الأصلي الذي اختيرت منه العينات يتبع توزيعاً طبيعياً ونذكرها بدون برهان.

نظرية (٧-٢) إذا أخذنا عينات متكررة من مجتمع معروف أنه يتبع توزيعاً طبيعياً بمتوسط μ وانحراف معياري σ فإن التوزيع العيني للإحصاء \bar{X} يتبع توزيعاً طبيعياً بمتوسط $\mu_{\bar{X}} = \mu$ وانحراف معياري $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ حيث $\mu_{\bar{X}}$ و $\sigma_{\bar{X}}$ يرمزان للمتوسط والانحراف المعياري على التوالي للتوزيع العيني للإحصاء \bar{X} .

مثال (٧-١) إذا كانت أوزان الطلاب في جامعة ما تتبع توزيعاً طبيعياً بمتوسط 70 كيلو جراماً وانحرافه المعياري 10 كيلو جراماً. اختيرت عينة عشوائية مكونة من 25 طالباً فما هو

احتمال أن يكون متوسط الأوزان أقل من 75 كيلو جرام.
الحل .

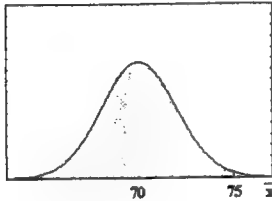
$\mu = 70$, $\sigma = 10$, $n = 25$
تبعا لنظرية (١-٧) فإن \bar{X} يتبع توزيعاً طبيعياً متوسطة $\mu_{\bar{X}} = \mu = 70$ وانحرافه المعياري :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2.$$

والمطلوب هو حساب الاحتمال :

$$P(\bar{X} < 75) .$$

والذي يساوي المساحة المظلة في شكل (١-٧) .



شكل (١-٧)

عندما $\bar{x} = 75$ فإن قيمة z المقابلة لها هي :

$$z = \frac{\bar{x} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{75 - 70}{2} = 2.5$$

وعلى ذلك :

$$\begin{aligned} P(\bar{X} < 75) &= P(Z < 2.5) \\ &= 1 - P(Z > 2.5) \\ &= 0.5 + P(0 < Z < 2.5) \\ &= 0.5 + 0.4938 = 0.9938 . \end{aligned}$$

في كثير من الحالات يكون التوزيع الاحتمالي للمجتمع الأصلي غير طبيعي ويتطلب الأمر معرفة التوزيع الاحتمالي للوسط الحسابي \bar{X} . للتسهيل سوف نستخدم مجتمعا منفصلا

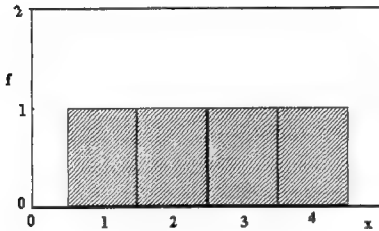
منتظما يتكون من القيم 1, 2, 3, 4 والذي متوسطه :

$$\mu = \frac{1+2+3+4}{4} = 2.5$$

والمحراف المعياري :

$$\sigma = \sqrt{\frac{(1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-2.5)^2}{4}} = 1.118.$$

المرج التكراري لهذا المجتمع موضح في شكل (٧-٢).



شكل (٧-٢)

بفرض أنه تم اختيار كل العينات من الحجم $n = 2$ من هذا المجتمع بإرجاع والسؤال يكافئ المعاينة من مجتمع لا نهائي. يعطى جدول (٧-١) كل العينات الممكنة التي يمكن اختيارها (عدد العينات $N^n = 4^2 = 16$ عينة) من هذا المجتمع مع قيمها. لكل عينة تم حساب \bar{x} والتوزيع التكراري لمجتمع متوسط العينات التي حجم كل منها $n = 2$ معطى في جدول (٧-٢) وتوزيعه التكراري في شكل (٧-٣).

جدول (٧-١)

رقم العينة	1	2	3	4	5	6	7	8
القيم	1,1	1,2	1,3	1,4	2,1	2,2	2,3	2,4
رقم العينة	9	10	11	12	13	14	15	16
القيم	3,1	3,2	3,3	3,4	4,1	4,2	4,3	4,4

جدول (٧-٢)

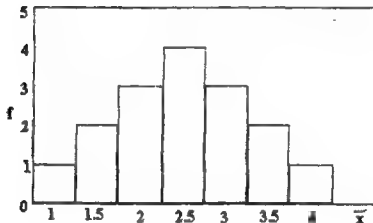
\bar{x}	1	1.5	2	2.5	3	3.5	4
f التكرار	1	2	3	4	3	2	1

يلاحظ أن توزيع المعاينة للإحصاء \bar{X} في شكل (٧-٣) تقريبا طبيعي. المتوسط والانحراف المعياري للتوزيع العيني للإحصاء \bar{X} تم حسابهما من جدول (٧-٢) وهما على التوالي :

$$\mu_{\bar{X}} = \frac{\sum f \bar{x}}{\sum f} = \frac{40}{16} = 2.5 = \mu ,$$

$$\sigma_{\bar{X}} = \sqrt{\frac{\sum f (\bar{x} - 2.5)^2}{\sum f}} = \sqrt{\frac{10}{16}}$$

$$= 0.791 = \frac{1.118}{\sqrt{2}} = \frac{\sigma}{\sqrt{n}} .$$



شكل (٧-٣)

دائما المتوسط للإحصاء \bar{X} يساوي متوسط المجتمع الذي اختيرت منه العينات العشوائية ولا يعتمد على حجم العينة. بينما الانحراف المعياري للإحصاء \bar{X} يعتمد على حجم العينة ويساوي الانحراف المعياري للمجتمع الأصلي σ مقسوما على \sqrt{n} . وعلى ذلك كلما زادت حجم العينة كلما قل الخطأ المعياري للإحصاء \bar{X} واقتربت \bar{x} من μ وعلى ذلك يمكننا استخدام \bar{x} كتقدير للمعلمة μ .

نظرية (٧-٣) إذا اختيرت كل العينات الممكنة من الحجم n بإرجاع من مجتمع محدود

من الحجم N وله متوسط μ وانحراف معياري σ فإن التوزيع العيني للإحصاء \bar{X} تقريباً يتبع توزيعاً طبيعياً بمتوسط $\mu_{\bar{X}} = \mu$ وانحراف معياري $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ وعلى ذلك :

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

هي قيمة لمغير عشوائي Z يتبع التوزيع الطبيعي القياسي . نظرية (٧-٣) صحيحة لأي مجتمع محدود عندما $n \geq 30$.

مثال (٧-٢) مجتمع مكون من المفردات الآتية :

2 , 2 , 2 , 4 , 5 , 7 , 7 , 7 , 8

أوجد احتمال أن عينة عشوائية من الحجم $n = 35$ اختيرت من هذا المجتمع بإرجاع تعطى متوسط عينة أكبر من 5.

الحل . المتوسط والانحراف المعياري للمجتمع المعطى هما :

$$\mu = 4.889 , \quad \sigma = 2.331 .$$

وحيث أن $n > 30$ وبما لنظرية (٧-٣) فإن التوزيع العيني للإحصاء \bar{X} تقريباً يتبع التوزيع الطبيعي بمتوسط $\mu_{\bar{X}} = \mu = 4.889$ وانحراف

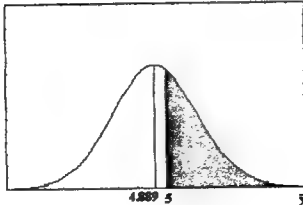
معيارى $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2.331}{\sqrt{35}} = 0.394$. احتمال أن \bar{X} أكبر من 5 يساوى المساحة

المظللة في شكل (٧-٤) . قيمة z المقابلة لقيمة $\bar{X} = 5$ هي :

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{5 - 4.889}{0.394} = 0.282 .$$

وعلى ذلك :

$$\begin{aligned} P(\bar{X} > 5) &= P(Z > 0.282) \\ &= 0.5 - P(0 < Z < 0.282) . \\ &= 0.5 - 0.1103 = 0.3897 . \end{aligned}$$



شكل (٧-٤)

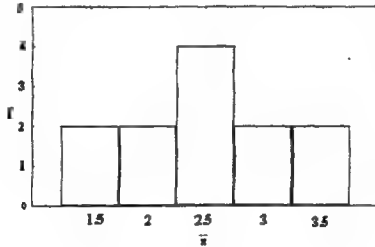
بفرض أننا سحبنا كل العينات الممكنة من الحجم $n = 2$ من مجتمعنا المنتظم والذي مشاهداته 1, 2, 3, 4 ولكن بدون إرجاع. لكل عينة تم حساب متوسط العينة \bar{x} . يعطى جدول (٣-٧) كل العينات الممكنة التي يمكن اختيارها من هذا المجتمع بدون إرجاع مع قيم كل عينة (عدد العينات $= 12 = \frac{4!}{2!} = \frac{N!}{(N-n)!}$ عينة). التوزيع التكراري لمجتمع متوسط العينات من الحجم $n = 2$ معطى في جدول (٤-٧). المدرج التكراري لمجتمع متوسط العينات موضح في شكل (٥-٧).

جدول (٣-٧)

رقم العينة	1	2	3	4	5	6
القيم	1,2	1,3	1,4	2,1	2,3	2,4
رقم العينة	7	8	9	10	11	12
القيم	3,1	3,2	3,4	4,1	4,2	4,3

جدول (٤-٧)

\bar{x}	1.5	2.0	2.5	3.0	3.5
f	2	2	4	2	2



شكل (٥-٧)

يتضح من شكل (٥-٧) أن التوزيع العيني للإحصاء \bar{x} في حالة السحب بدون إرجاع من مجتمع محدود بعيدا عن التوزيع الطبيعي حيث $n = 2$. من جدول (٤-٧) يمكن

حساب المتوسط والانحراف المعياري للإحصاء \bar{X} على التوالي :

$$\mu_{\bar{X}} = \frac{\sum f \bar{x}}{\sum f} = \frac{30}{12} = 2.5 = \mu ,$$

$$\sigma_{\bar{X}} = \sqrt{\frac{\sum f(\bar{x} - 2.5)^2}{\sum f}} = \sqrt{\frac{5}{12}} = 0.645$$

$$= \frac{1.118}{\sqrt{2}} \sqrt{\frac{4-2}{4-1}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = 0.645.$$

عندما يكون $n \geq 30$ يمكن تطبيق النظرية التالية :

نظرية (٧-٤) إذا اختبرت كل العينات الممكنة بدون إرجاع من مجتمع محدود من الحجم N وله متوسط μ وانحراف معياري σ ، فإن التوزيع العيني للإحصاء \bar{X} تقريبا يتبع توزيعا طبيعيا بمتوسط وانحراف معياري معطى كالتالي :

$$\mu_{\bar{X}} = \mu,$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}.$$

يسمى المقدار $\sqrt{\frac{N-n}{N-1}}$ معامل التصحيح. إذا كان حجم العينة صغيرا جدا بالنسبة لحجم

المتجمع فإن $(N-1) / (N-n)$ تكون قريبة من 1 ويمكن إسقاطها من المعادلة. وقد جرت العادة على إهمال هذا الحد عندما تكون $n < 0.05 N$.

مثال (٧-٣) من مجتمع مكون من القيم

$$2, 2, 2, 4, 5, 6, 7, 7, 7, 8$$

المطلوب : (أ) حساب متوسط المجتمع الأصلي μ وانحرافه المعياري σ .

(ب) حساب متوسط مجتمع متوسطات العينات $\mu_{\bar{X}}$ وانحرافه المعياري

$\sigma_{\bar{X}}$ عند $n=2$ (السحب بدون إرجاع).

الحل .

$$\mu = 5, \sigma = 2.24, N = 10, n = 2 \quad (أ)$$

$$\mu_{\bar{X}} = \mu = 5 \quad (ب)$$

وحيث أن :

$$0.05 N = (0.05) (10) = 0.5$$

أي أن $n = 2$ سوف تكون أكبر من $0.5 (0.05 N)$ وعلى ذلك لا يمكننا إهمال معامل

التصحيح في صيغة $\sigma_{\bar{X}}$. وعلى ذلك :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{2.24}{\sqrt{2}} \sqrt{\frac{10-2}{10-1}} = 1.4933.$$

مثال (٦-٤) افترض مجتمعا ما يتكون من 1000 عنصر له متوسط حسابي $\mu=15$ وانحراف معياري $\sigma=6$ أوجد :

(أ) المتوسط الحسابي والانحراف المعياري لتوزيع العينة للإحصاء \bar{X} عندما يكون

(السحب بدون إرجاع) $n = 36$

(ب) احتمال أن يقع متوسط العينة العشوائية من الحجم $n = 36$ بين 13، 16.

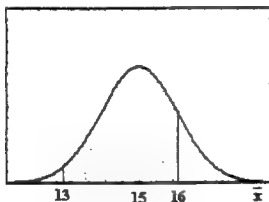
الحل .

(أ) بما أن $n = 36 < 0.05 N = 50$ فإن التوزيع العيني للإحصاء \bar{X} يكون له متوسط وانحراف معياري كالآتي :

$$\mu_{\bar{X}} = \mu = 15,$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{36}} = 1.$$

(ب) الاحتمال المطلوب موضح بالمساحة المظلمة في شكل (٦-٧) .



شكل (٦-٧)

قيمة z_1 المقابلة لقيمة $\bar{x}_1 = 13$ هي :

$$z_1 = \frac{\bar{x}_1 - \mu}{\sigma_{\bar{X}}} = \frac{13 - 15}{1} = -2,$$

قيمة z_2 المقابلة للقيمة $\bar{x}_2 = 16$ هي :

$$z_2 = \frac{\bar{x}_2 - \mu}{\sigma_{\bar{x}}} = \frac{16 - 15}{1} = 1.$$

وعلى ذلك :

$$\begin{aligned} P(13 < \bar{X} < 16) &= P(-2 < Z < 1) \\ &= P(0 < Z < 2) + P(0 < Z < 1) \\ &= 0.4772 + 0.3413 = 0.8185. \end{aligned}$$

للمجتمعات الكبيرة أو اللانهائية سواء كانت متصلة أم مقطعة تنص النظرية التالية على .
نظرية (٥-٧) إذا اختبرت كل العينات الممكنة من الحجم n من مجتمع كبير أو لانهائي
بمتوسط μ وتباين σ^2 فإن التوزيع العيني للإحصاء \bar{X} تقريبا يتبع توزيعها طبيعيا بمتوسط :

$$\mu_{\bar{X}} = \mu$$

وانحراف معياري :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}.$$

وعلى ذلك فإن :

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

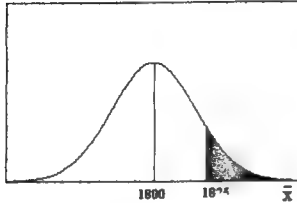
هي قيمة لمغير عشوائي Z يتبع التوزيع الطبيعي القياسي .

التقريب الطبيعي في نظرية (٥-٧) سوف يكون جيدا إذا كانت $n \geq 30$ بصرف
النظر عن شكل المجتمع الأصلي الذي اختبرت منه العينات. إذا كانت $n < 30$ التقريب
سوف يكون جيد فقط إذا كان المجتمع لا يختلف كثيرا عن التوزيع الطبيعي .
مثال (٥-٧) إذا كانت أعمار المصايح المنتجة بواسطة أحد المصانع لها متوسط عمر
 $\mu = 1800$ ساعة وانحراف معياري $\sigma = 200$ ساعة. أوجد احتمال أن عينة عشوائية من
100 مصباح سوف يكون لها متوسط عمر أكبر من 1825 ساعة.
الحل .

المجتمع كبير والعينة هنا كبيرة $n=100$ ، على ذلك التوزيع العيني للإحصاء \bar{X} تقريبا
يتبع توزيعها طبيعيا بمتوسط 1800 ساعة وانحراف معياري $= 20 = \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{100}}$. وعلى
ذلك عندما يكون $\bar{x} = 1825$ فإن :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{1825 - 1800}{20} = 1.25.$$

الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧-٧).



شكل (٧-٧)

وعلى ذلك :

$$P(\bar{X} > 1825) = P(Z > 1.25) \\ = 0.5 - 0.3944 = 0.1056.$$

مثال (٧-٦) إذا كان متوسط الدب الأسترالي $\mu = 20$ بوصة بانحراف معياري $\sigma = 4$ بوصة. فإذا خطط لاختبار عينة عشوائية من الحجم $n = 64$. ما هو احتمال أن يكون متوسط العينة أكبر من 21 بوصة.

الحل . التوزيع العيني للإحصاء \bar{X} تقريبا يتبع توزيعها طبيعيا بمتوسط $\mu = 20$ وانحراف معياري :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{64}} = 0.5.$$

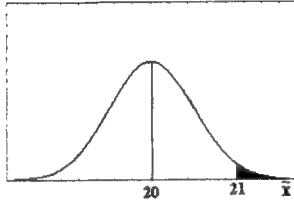
احتمال أن يكون متوسط عينة عشوائية من الحجم $n = 64$ أكبر من 21 بوصة يساوى المساحة المظللة في شكل (٧-٨).

عندما يكون $\bar{x} = 21$ فإن :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{X}}} = \frac{21 - 20}{0.5} = 2.$$

وعلى ذلك :

$$P(\bar{X} > 21) = P(Z > 2) = 0.5 - P(0 < Z < 2) \\ = 0.5 - 0.4772 = 0.0228.$$



شكل (٧-٨)

مثال (٧-٧) إذا كان متوسط سمك الرخام المنتج في مصنع للعب الأطفال هو 0.85 سم باحتراف معياري 0.01 المطلوب :

(أ) احتمال اختيار عينة عشوائية من 100 قطعة رخام لها سمك أكبر من 0.851.

(ب) ما هما القيمتان التي نتوقع أن يكون 95% من متوسطات العينات بينهما .

الحل .

(أ) التوزيع المعنى للإحصاء \bar{X} تقريبا يتبع توزيعا طبيعيا بمتوسط $\mu_{\bar{X}} = \mu = 0.85$ وانحراف معياري :

$$\frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{100}} = 0.001.$$

عندما يكون $\bar{x} = 0.851$ فإن :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{X}}} = \frac{0.851 - 0.85}{0.001} = 1.0.$$

الاحتمال المطلوب يساوي المساحة المظلة في شكل (٧-٩) وعلى ذلك :

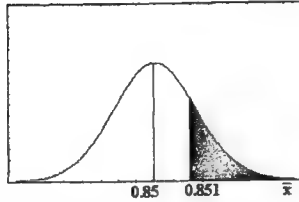
$$P(\bar{X} > 0.851) = P(Z > 1.0) \\ = 0.5 - 0.3413 = 0.1587.$$

(ب) باستخدام القاعدة التجريبية فإننا نتوقع أن 95% متوسطات العينات تقع في

الفترة :

$$\mu_{\bar{X}} \pm 2\sigma_{\bar{X}} = 0.85 \pm 2(0.001)$$

أي في الفترة (0.848, 0.852) .



شكل (٩-٧)

(٤-٧) التوزيعات العينية للفرق بين متوسطي مجتمعين

Sampling Distributions of the Different Between Two Populations Means

بفرض أن لدينا مجتمعين الأول متوسطه μ_1 وتباينه σ_1^2 والثاني متوسطه μ_2 وتباينه σ_2^2 . بفرض أن قيم المتغير \bar{X}_1 تمثل متوسطات لعينات عشوائية من الحجم n_1 اختبرت من المجتمع الأول، وقيم المتغير \bar{X}_2 تمثل متوسطات لعينات عشوائية من المجتمع الثاني ومستقلة عن المجتمع الأول. التوزيع للفرق $\bar{X}_1 - \bar{X}_2$ بين القمتين من متوسطات العينتين المستقلتين يسمى التوزيع العيني للإحصاء $\bar{X}_1 - \bar{X}_2$. للتسهيل نفرض أن المجتمع الأول من الحجم $N_1 = 3$ ويتكون من القيم 4, 5, 6 والتي متوسطها :

$$\mu_1 = \frac{4 + 5 + 6}{3} = 5$$

وتباينها :

$$\sigma_1^2 = \frac{(4-5)^2 + (5-5)^2 + (6-5)^2}{3} = \frac{2}{3}$$

المجتمع الثاني يتكون من القيمتين 1, 4 ولهما المتوسط :

$$\mu_2 = \frac{1 + 4}{2} = 2.5$$

والتباين :

$$\sigma_2^2 = \frac{(1-2.5)^2 + (4-2.5)^2}{2} = \frac{9}{4}.$$

من المجتمع الأول تم اختيار كل العينات الممكنة من الحجم $n_1 = 2$ مع الإرجاع وحساب المتوسط \bar{x} لكل عينة. بنفس الشكل للمجتمع الثاني تم اختيار كل العينات الممكنة من الحجم $n_2 = 3$ وحساب \bar{x}_2 لكل عينة. الفئتان من كل العينات ومتوسطاتها معطاة في جدول (٥-٧).

جدول (٥-٧)

المجتمع الأول	رقم العينة	1	2	3	4	5	6	7	8	9
	القيم	4,4	4,5	4,6	5,4	5,5	5,6	6,4	6,5	6,6
	\bar{x}_1	4.0	4.5	5.0	4.5	5.0	5.5	5.0	5.5	6.0
المجتمع الثاني	رقم العينة	1	2	3	4	5	6	7	8	
	القيم	1,1,1	1,1,4	1,4,1	4,1,1	4,4,1	1,4,4	4,1,4	4,4,4	
	\bar{x}_2	1	2	2	2	3	3	3	4	

الفروق الممكنة والتي عددها 72 من $\bar{x}_1 - \bar{x}_2$ معطاة في جدول (٦-٧).

جدول (٦-٧)

\bar{x}_2	\bar{x}_1								
	4.0	4.5	5.0	4.5	5.0	5.5	5.0	5.5	6.0
1	3.0	3.5	4.0	3.5	4.0	4.5	4.0	4.5	5.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
4	0.0	0.5	1.0	0.5	1.0	1.5	1.0	1.5	2.0

التوزيع التكراري للإحصاء $\bar{x}_1 - \bar{x}_2$ في جدول (٧-٧) ومدرجة التكراري في شكل)

(١٠-٧). من الواضح أن المنظر العشوائي $\bar{x}_1 - \bar{x}_2$ تقريبا يتبع توزيعا طبيعيا وهذا

التقريب يتحسن عندما يزيد n_1, n_2 . بتطبيق نظرية (٤-٦) ثم نظرية (٧-٣) فإن متوسط الفروق لجمع العينات المستقلة هو :

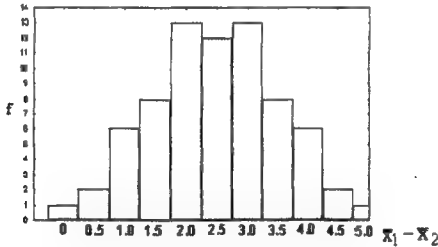
$$\mu_{\bar{X}_1 - \bar{X}_2} = E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2.$$

هذه النتيجة يمكن التحقق منها من البيانات في جدول (٧-٧) حيث :

$$\begin{aligned}\mu_{\bar{X}_1 - \bar{X}_2} &= \frac{\sum f(\bar{X}_1 - \bar{X}_2)}{\sum f} = \frac{180}{72} = 2.5 = 5 - 2.5 \\ &= \mu_1 - \mu_2.\end{aligned}$$

جدول (٧-٧)

$\bar{X}_1 - \bar{X}_2$	0.0	.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
f	1	2	6	8	13	12	13	8	6	2	1



شكل (٧-١٠)

أيضا بتطبيق نظرية (٤-٨) ثم نظرية (٧-٣) فإن التباين لفروق المتوسطات المستقلة هو :

$$\begin{aligned}\sigma_{\bar{X}_1 - \bar{X}_2}^2 &= \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \\ &= \left[\left(\frac{2}{3} \right) / 2 \right] + \left[\left(\frac{9}{4} \right) / 3 \right] = 1.08333.\end{aligned}$$

هذه النتيجة يمكن التحقق منها بسهولة وذلك بحساب التباين (1.08333) من البيانات في جدول (٧-٧). النتائج التي تم الحصول عليها للتوزيع العيني للإحصاء $\bar{X}_1 - \bar{X}_2$ عند المعاينة بإرجاع من مجتمع محدود تكون صحيحة للمجتمعات اللانهائية سواء المتقطعة أو المتصلة وأيضا للمجتمعات المحدودة عند المعاينة بدون إرجاع بشرط أن أحجام المجتمعات N_1, N_2 تكون كبيرة نسبيا عن أحجام العينات n_1, n_2 على التوالي. أما إذا كان حجم المجتمع صغيرا والسحب بدون إرجاع فلا بد من حساب $\sigma_{\bar{X}_1}^2, \sigma_{\bar{X}_2}^2$ من صيغة \bar{X} في نظرية (٧-٤).

تقتصر الدراسة من الآن وفي الفصول التالية على التوزيع العيني للفروق بين المتوسطات المستقلة فقط إذا كان حجم المجتمع الذي تختار منه العينات كبيراً أو لانهائياً. نظرية (٧-٦) إذا اختبرت عينات مستقلة من الحجم N_1, N_2 من مجتمعين كبيرين (أو لانهائيتين)، مقطعة أو متصلة، بمتوسطي μ_1, μ_2 وتبايني σ_1^2, σ_2^2 على التوالي، فإن التوزيع العيني لفروق المتوسطات، $\bar{X}_1 - \bar{X}_2$ ، تقريباً يتبع توزيعاً طبيعياً بمتوسط والانحراف المعياري معطى كالتالي:

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2,$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

وعلى ذلك :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

هي قيمة لمغير عشوائي Z يتبع التوزيع الطبيعي القياسي.

إذا كان كل من n_1, n_2 أكبر من أو يساوي 30، فإن التقريب الطبيعي لتوزيع $\bar{X}_1 - \bar{X}_2$ يكون جيداً جداً.

ملحوظة : إذا كانت العينتين تم اختيارهما من مجتمعين طبيعيين فإن :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

هي قيمة لمطير عشوائي Z يتبع التوزيع الطبيعي القياسي بصرف النظر عن حجم كلا من n_1, n_2

مثال (٧-٨) ينتج المصنع A بطاريات سيارة لها متوسط عمر 3.5 سنة بانحراف معياري 0.45 سنة. نفس البطاريات تنتج من المصنع B بمتوسط عمر 3.3 سنة وانحراف معياري 0.3 سنة. ما هو احتمال أن عينة عشوائية من 30 بطارية من المصنع A يكون لها متوسط عمر على الأقل يزيد 0.4 سنة عن متوسط عمر 36 بطارية من المصنع B ؟
الحل . سوف يكون لدينا البيانات التالية :

المجتمع الأول	المجتمع الثاني
$\mu_1 = 3.5$	$\mu_2 = 3.3$
$\sigma_1 = 0.45$	$\sigma_2 = 0.3$
$n_1 = 30$	$n_2 = 36$

العينات هنا كبيرة بدرجة كافية بحيث أن كل من \bar{X}_1 و \bar{X}_2 تقريبا يتبع توزيعها طبيعيًا وعلى ذلك فإن $\bar{X}_1 - \bar{X}_2$ تقريبا تتبع توزيعها طبيعيًا بمتوسط وانحراف معياري على التوالي :

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2 = 3.5 - 3.3 = 0.2,$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{0.45^2}{30} + \frac{0.3^2}{36}} = 0.096177,$$

الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧-١١).

$$\bar{X}_1 - \bar{X}_2 = 0.4 \quad \text{عندما :}$$

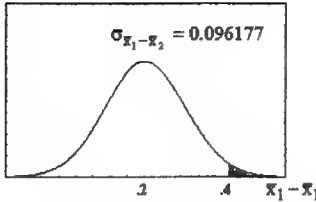
لأن :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{0.4 - 0.2}{0.096177} = 2.08 ,$$

وعلى ذلك :

$$\begin{aligned} P(\bar{X}_1 - \bar{X}_2 > 0.4) &= P(Z > 2.08) \\ &= 0.5 - P(0 < Z < 2.08) \end{aligned}$$

$$= 0.5 - 0.4812 = 0.0188.$$



شكل (٧-١١)

(٧-٥) التوزيعات العينية للنسب

Sampling Distributions of Proportions

بفرض أن لدينا مجتمعا ما وأن بعض مفردات هذا المجتمع تتوفر فيها صفة معينة وأن نسبة هذه المفردات هي p . فعلى سبيل المثال p قد تكون نسبة الأفراد المصابين بتسوس الأسنان في مدينة ما أو نسبة الطلبة المدخنين في كلية ما أو نسبة الوحدات المعيبة في مصنع ما... الخ. إذا أخذنا عينة عشوائية من الحجم n من هذا المجتمع ووجدنا من بينها x مفردة تتوفر فيها الصفة، وتم حساب $\hat{p} = \frac{x}{n}$ والتي تمثل نسبة المفردات في العينة والتي تتوفر فيها الصفة المعنية. إذا أخذنا عينات متكررة من الحجم n من هذا المجتمع فإن \hat{p} تتغير من عينة إلى أخرى وتمثل قيمة للإحصاء \hat{P} . الآن سوف نتعرف على كيفية اشتقاق التوزيع العيني للنسبة وخصائصه سواء في حالة السحب بإرجاع أو بدون إرجاع وذلك من الأمثلة التالية:

مثال (٧-٩) مجتمع يتكون من القيم 1, 2, 3, 4 فإذا تم سحب كل العينات الممكنة من الحجم $n = 2$ من هذا المجتمع (إرجاع). المطلوب إيجاد التوزيع العيني للإحصاء \hat{P} والذي يمثل نسبة ظهور الرقم 4 في العينة. وإثبات أن المتوسط والتباين للتوزيع العيني للإحصاء \hat{P} هما على التوالي :

$$\mu_{\hat{p}} = p,$$

$$\sigma_{\hat{p}}^2 = \frac{pq}{n}.$$

الحل . الجدول (٧-٨) يحتوي على كل العينات الممكنة ونسبة ظهور الرقم 4 فيها.

جدول (٧-٨)

رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4	رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4
1	1,1	0	0	9	3,1	0	0
2	1,2	0	0	10	3,2	0	0
3	1,3	0	0	11	3,3	0	0
4	1,4	1	0.5	12	3,4	1	.5
5	2,1	0	0	13	4,1	1	.5
6	2,2	0	0	14	4,2	1	.5
7	2,3	0	0	15	4,3	1	.5
8	2,4	1	0.5	16	4,4	2	1

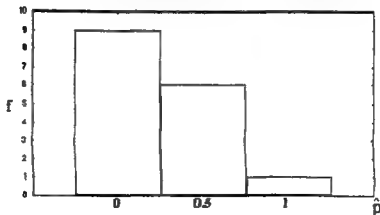
التوزيع التكراري لنسبة ظهور الرقم 4 للعينات من الحجم $n = 2$ التي تم اختيارها من المجتمع الذي حجمه $N = 4$ (إرجاع) معطاة في جدول (٧-٩) ومدرجها التكراري في شكل (٧-١٢).

جدول (٧-٩)

\hat{p}	0	0.5	1
f	9	6	1

التوزيع التكراري في شكل (٧-١٢) ملئو ناحية اليمين وذلك لأن $p < 0.5$. إذا كانت $p > 0.5$ فإن التوزيع سوف يكون ملئو ناحية اليسار .

شكل (٧-١٢)



الموسط والتباين للتوزيع العيني للإحصاء \hat{P} تم حسابهما من جدول (٧-٩) وهما :

$$\mu_{\hat{P}} = \frac{\sum f \hat{p}}{\sum f} = \frac{4}{16} = 0.25 = p,$$

$$\begin{aligned} \sigma_{\hat{P}}^2 &= \frac{\sum f (\hat{p} - 0.25)^2}{\sum f} = 0.094 \\ &= \frac{(0.25)(0.75)}{2} = \frac{pq}{n}. \end{aligned}$$

مثال (٧-١٠) أوجد التوزيع العيني للإحصاء \hat{P} للبيانات في مثال (٧-٩) إذا كان السحب بدون إرجاع وأثبت أن موسط وتباين التوزيع العيني للإحصاء \hat{P} هما على التوالي :

$$\mu_{\hat{P}} = p,$$

$$\sigma_{\hat{P}}^2 = \frac{pq}{n} \left(\frac{N-n}{N-1} \right).$$

الجدول (٧-١٠) يحتوى على كل العينات ونسبة ظهور الرقم 4 فيها .

جدول (٧-١٠)

رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4	رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4
1	1,2	0	0	7	3,1	0	0
2	1,3	0	0	8	3,2	0	0
3	1,4	1	.5	9	3,4	1	.5
4	2,1	0	0	10	4,1	1	.5
5	2,3	0	0	11	4,2	1	.5
6	2,4	1	.5	12	4,3	1	.5

التوزيع التكراري لنسبة ظهور الرقم 4 للعينات من الحجم $n = 2$ والتي تم اختيارها من المجتمع الذى حجمه $N = 4$ (بدون إرجاع) معطاة في جدول (٧-١١) ومدرجها التكراري في شكل (٧-١٣).

جدول (١١-٧)

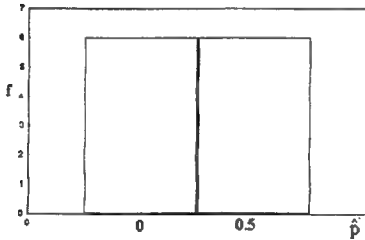
\hat{p}	0	0.5
f	6	6

المتوسط والتباين للتوزيع العيني للإحصاء \hat{P} تم حسابهما من جدول (١١-٧) وهما :

$$\mu_{\hat{p}} = \frac{\sum f \hat{p}}{\sum f} = \frac{3}{12} = .25 = p,$$

$$\sigma_{\hat{p}}^2 = \frac{\sum f (\hat{p} - 0.25)^2}{\sum f} = 0.0625.$$

$$= \frac{(0.25)(0.75)}{2} \left(\frac{4-2}{4-1} \right) = \frac{pq}{n} \left(\frac{N-n}{N-1} \right)$$



شكل (١٣-٧)

يتضح من المثال السابق أن معامل التصحيح $\left(\frac{N-n}{N-1} \right)$ في صيغة $\sigma_{\hat{p}}^2$ يستعمل إذا كان المجتمع محدود والسحب بدون إرجاع. إذا كان حجم العينة أصغر من $0.05N$ يمكن اعتبار

$$\left(\frac{N-n}{N-1} \right) = 1$$

مثال (١١-٧) إذا كانت نسبة المدخنين في مجتمع من الحجم $N=350$ هو 0.4.

سحبت عينة عشوائية من الحجم $n = 60$ من هذا المجتمع (بدون إرجاع) أوجد المتوسط والتباين للتوزيع العيني للإحصاء \hat{P} .
الحل .

$$\mu_{\hat{P}} = .4,$$

التباين بحسب من الصيغة :

$$\sigma_{\hat{P}}^2 = \frac{pq}{n} \left(\frac{N-n}{N-1} \right).$$

وذلك لأن $n=60 > 0.05N=17.1$ • وعلى ذلك :

$$\sigma_{\hat{P}}^2 = \frac{pq}{n} \left(\frac{N-n}{N-1} \right) = \frac{(0.4)(0.6)}{60} \left(\frac{350-60}{350-1} \right) = 0.0033.$$

للمجتمعات الكبيرة أو اللانهائية فإن التوزيع العيني للإحصاء \hat{P} تحدده النظرية التالية :
نظرية (٧-٧) إذا كانت p هي نسبة صفة معينة في مجتمع ما واختيرت من هذا المجتمع عينات كبيرة، حجم كل منها n وكان الإحصاء \hat{P} يمثل نسبة وجود هذه الصفة في العينات فإن \hat{P} تقريباً تتبع توزيعاً طبيعياً متوسطة وتباينه على التوالي :

$$\mu_{\hat{P}} = p,$$

$$\sigma_{\hat{P}}^2 = \frac{pq}{n}.$$

وعلى ذلك :

$$z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

هي قيمة لمطير عشوائى Z تقريباً يتبع التوزيع الطبيعي القياسى .
وضع العالم (1963) Cochran قواعد لحجم العينة اللازم لتطبيق نظرية (٧-٧)
(٧) معطاة في جدول (٧-١٢) .

جدول (٧-١٢)

إذا كانت p تساوى	يستخدم التقريب الطبيعي إذا كان n على الأقل يساوى
.5	30
.4 - .6	50
.3 - .7	80
.2 - .8	200
.1 - .9	600
.05 - .95	1400

مثال (٧-١٢) بفرض أن مجتمعاً ما أفرادهُ عدة آلاف يمثل مصنعاً لإنتاج كسرات المعاييدة. فإذا كان 0.2 من الكسرات المنتجة تالفة (أ) أوجد المتوسط والانحراف للتوزيع العيني للإحصاء \hat{P} وذلك عندما يكون $n = 300$ (ب) ما هو الاحتمال أن عينة عشوائية من الحجم $n = 300$ تعطى نسبة صفة أكبر من 0.19 (ج) أوجد احتمال أن يكون نسبة الكسرات المنتجة التالفة تزيد عن 0.052 .
الحل .

(أ) بما أن المجتمع كبيراً و $p = 0.2$, $n = 300$ تحققان القواعد في جدول (٧-١٢) وعلى ذلك فإن التوزيع العيني للإحصاء \hat{P} تقريباً يتبع توزيعاً طبيعياً بمتوسط:

$$\mu_{\hat{P}} = p = 0.2$$

وانحراف معياري :

$$\sigma_{\hat{P}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.2)(0.8)}{300}} = 0.023.$$

(ب) عندما يكون $\hat{p} = 0.19$ فإن

$$z = \frac{\hat{p} - p}{\sigma_{\hat{P}}} = \frac{0.19 - 0.2}{0.023} = -0.43.$$

وعلى ذلك فإن :

$$\begin{aligned} P(\hat{P} > 0.19) &= P(Z > -0.43) \\ &= 0.5 + p(0 < Z < 0.43) \\ &= 0.5 + 0.1664 = 0.6664. \end{aligned}$$

(ج) عندما يكون $\hat{p} = 0.052$ فإن :

$$z = \frac{0.052 - 0.2}{0.023} = -6.43,$$

$$P(\hat{P} > 0.052) = P(Z > -6.43) \approx 1$$

مثال (١٣-٧) للمثال (١١-٧) أوجد احتمال أن تكون نسبة المدخنين في العينة أكبر من 0.35 .

الحل . عندما يكون $\hat{p} = 0.35$ فإن :

$$z = \frac{0.35 - 0.4}{0.058} = -0.86$$

وعلى ذلك فإن :

$$\begin{aligned} P(\hat{P} > 0.35) &= P(Z > -0.86) \\ &= 0.5 + P(0 < Z < 0.86) \\ &= 0.5 + 0.3051 \\ &= 0.8051. \end{aligned}$$

يمكن تطبيق نظرية (٧-٧) في حالة السحب بدون إرجاع من مجتمع محدود إذا كان حجم العينة أكبر من $0.05N$ حيث متوسط التوزيع العيني للإحصاء \hat{P} في هذه الحالة مازال $\mu_{\hat{p}} = p$ ولكن التباين يساوى :

$$\sigma_{\hat{P}}^2 = \frac{pq}{n} \left(\frac{N-n}{N-1} \right).$$

وعلى ذلك فإن :

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}}$$

هي قيمة لمطير عشوائي Z يتبع التوزيع الطبيعي القياسي.

مثال (١٤-٧) إذا كانت نسبة المصابين بتسوس الأسنان في مجتمع من الحجم $N=200$ هو 0.3. سحبت عينة عشوائية من الحجم $n=80$ بدون إرجاع أوجد احتمال أن تكون نسبة المصابين بتسوس الأسنان في العينة أكبر من 0.35.

الحل . حجم العينة ($n=80$) أكبر من $0.05N=10$. التوزيع العيني للإحصاء \hat{P} سوف يكون تقريباً يتبع التوزيع الطبيعي لأن $n=80$ تحقق القواعد في جدول (١٢-٧). المتوسط و الانحراف المعياري للإحصاء \hat{P} سوف يكون :

$$\mu_{\hat{p}} = p = 0.3 ,$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}}$$

$$= \sqrt{\frac{(.3)(.7)}{80}} \sqrt{\frac{200-80}{199}} = 0.03979.$$

عندما يكون $\hat{p} = 0.35$ فإن :

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0.35 - 0.3}{0.03979} \approx 1.26.$$

$$P(\hat{P} > .35) = P(Z > 1.26) = 0.5 - P(0 < Z < 1.26)$$

$$= 0.5 - 0.3962 = 0.1038.$$

إذا كان لدينا مجتمعان مستقلان (مجتمعات كبيرة أو لانهائية) وإذا كانت p_1 هي نسبة توفر صفة ما في المجتمع الأول وكانت p_2 هي نسبة توفر الصفة نفسها في المجتمع الثاني . إذا اخترنا عينة عشوائية كبيرة حجمها n_1 من المجتمع الأول وحسبنا منها نسبة توفر الصفة محل الدراسة ولتكن \hat{p}_1 . وإذا اخترنا عينة عشوائية كبيرة حجمها n_2 من المجتمع الثاني وحسبنا منها نسبة توفر الصفة المطلوبة ولتكن \hat{p}_2 . بتكرار المعاينة من الحجم n_1 و n_2 فإن التوزيع العيني للإحصاء $\hat{P}_1 - \hat{P}_2$ تحدده النظرية التالية :

نظرية (٧-٨) التوزيع العيني للإحصاء $\hat{P}_1 - \hat{P}_2$ تقريباً يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{P}_1 - \hat{P}_2} = p_1 - p_2,$$

وتباين :

$$\sigma_{\hat{P}}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}.$$

وعلى ذلك تكون :

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

هي قيمة لتغير عشوائي Z الذي يتبع التوزيع الطبيعي القياسي .

تعطى النظرية (٧-٨) نتائجاً جيدة ، إذا كانت n_1 , n_2 محدداً طبقاً لقواعد Cochran المعطاة في جدول (٧-١٢) .

t Distribution

(٧-٦) توزيع t

في معظم الأبحاث وغالباً يكون تباين المجتمع الذي تختار منه العينات مجهولاً . للعينات العشوائية من الحجم $n \geq 30$ فإن التقدير الجيد للمعلمة σ^2 هو s^2 . إذا كانت $n \geq 30$

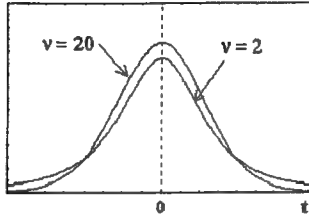
واستبدلنا σ بالقيمة S في صيغة Z لنظرية (٧-٥) فإن :

$$Z = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

هي قيمة لمغير عشوائي Z تقريبا يتبع التوزيع الطبيعي القياسي . أما إذا كان حجم العينة صغير ($n < 30$) فإن قيم $(\bar{X} - \mu)/(s/\sqrt{n})$ لا تتبع التوزيع الطبيعي القياسي . في هذه الحالة يكون احتمالا بتوزيع لإحصاء ما سوف نرمز له بالرمز T ، والذي قيمه تعطى من الصيغة التالية :

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

لقد تمكن ستودنت "Student" وهو لقب لعالم إحصائي، كان ينشر أبحاثه بتوقيع ستودنت، أن يشتق العبارة المضبوطة لتوزيع t ويسمى هذا التوزيع في كتب الإحصاء المختلفة "توزيع t " أو "توزيع t " . يشبه توزيع t التوزيع الطبيعي القياسي فكلاهما متماثل حول الصفر كما أن كلا التوزيعين هما شكل الناقوس ولكن توزيع t أكثر تشتتا وذلك راجع إلى الحقيقة أن قيم t تعتمد على الاختلاف في قيمتي \bar{X} و s^2 بينما قيم z تعتمد فقط على التغير في قيمة \bar{X} من عينة إلى أخرى . يختلف توزيع المتغير T عن المتغير Z في أن التباين يعتمد على حجم العينة n ودائما أكبر من الواحد الصحيح ، فقط عندما $n \rightarrow \infty$ فإن التوزيعين يتساويان. المقام $(n-1)$ والذي يظهر في صيغة s^2 يسمى درجات الحرية degree of freedom المرتبط بتباين العينة s^2 . بتكرار المعاينة من الحجم n وحساب \bar{X} و s^2 لكل عينة، فإن قيم t المقابلة يقال أنها تتبع توزيع t بدرجات حرية v ، حيث $v = n-1$. وعلى ذلك سوف يكون لدينا منحنيات t مختلفة أو توزيع t لكل حجم عينة . من خصائص توزيع t أنه كلما كبرت درجات الحرية v زاد ارتفاع منحنى t وأصبح أكثر تدبنا أي أقل تشتتا وفي النهاية ينطبق على منحنى التوزيع الطبيعي القياسي . المنحنى في شكل (٧-١) بدرجات حرية $v = 2$ يمثل توزيع كل قيم t المحسوبة من عينات عشوائية من الحجم $n = 3$ تكرر اختيارها من مجتمع طبيعي . بنفس الشكل، المنحنى بدرجات حرية $v = 20$ يمثل توزيع كل قيم t المحسوبة من عينات من الحجم $n = 21$.



شكل (٧-١٤)

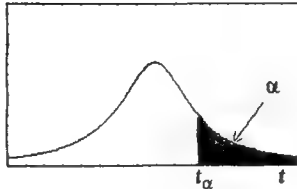
نظرية (٧-٩) إذا كان \bar{x} و s^2 هما المتوسط الحسابي والتباين على التوالي لعينة عشوائية من الحجم n مأخوذة من مجتمع طبيعي له متوسط μ وتباين σ^2 غير معروف فإن :

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

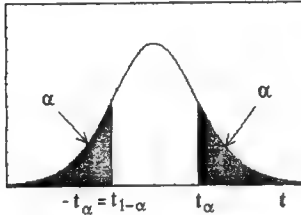
هي قيمة لمتغير عشوائي T له توزيع t بـ $v = n - 1$ درجات حرية .

بفرض أن t_α ترمز لقيمة t التي توجد على المحور الأفقي تحت منحنى توزيع t بـ v درجات حرية والتي المساحة على يمينها قدرها α كما هو موضح في شكل (٧-١٥).

شكل (٧-١٥)



الجدول في ملحق (٤) يعطى قيم t_{α} التي تتناظر الاحتمال α للدرجات حرية v حيث α تأخذ القيم التالية : 0.0005, 0.001, 0.005, 0.01, 0.025, 0.05, 0.10. ودرجات الحرية تأخذ القيم من $v=1$ إلى $v=\infty$. يوضح الصف الثاني من الجدول قيم α والعمود الأول من الشمال قيم درجات الحرية v . أما محتويات الجدول فهي القيم t_{α} . ولأن المنحنى متماثل فإن $t_{1-\alpha} = -t_{\alpha}$ كما هو موضح في شكل (١٦-٧).



شكل (١٦-٧)

مثال (١٥-٧) أوجد (أ) قيمة $t_{0.005}$, $v = 15$.

(ب) قيمة $t_{0.995}$, $v = 15$

الحل .

(أ) بالبحث في جدول توزيع t في ملحق (٤) عند تقاطع الصف $v = 15$

والعمود $\alpha = 0.005$ نجد أن $t_{0.005} = 2.947$.

(ب) باستخدام خاصية التماثل لمنحنى توزيع t فإن

$$t_{0.995} = -t_{0.005} \text{ أى أن } t_{0.995} = -2.947$$

مثال (١٦-٧) أوجد قيمة α حيث

$$v = 16, t_{\alpha} = -1.746$$

الحل .

حيث أن قيمة t سالبة فإنها تقع في الذيل الأيسر من توزيع t وباستخدام خاصية

التماثل لمنحنى توزيع t فإن :

$$t_{1-\alpha} = -t_{\alpha} = 1.746$$

ومن جدول توزيع t في ملحق (٤) فإن $1-\alpha = 0.05$ ومنها $\alpha = 0.95$.
 مثال (٧-١٧) إذا كانت أعمار المصاييح الكهربائية المنتجة بواسطة أحد المصانع تتبع توزيعاً طبيعياً ، ويدعي صاحب المصنع أن متوسط أعمار هذه المصاييح هو $\mu = 500$ ساعة . لمتابعة جودة الإنتاج يأخذ 20 مصباحاً كل شهر ، ويحقق الإنتاج للمواصفات القياسية إذا كانت قيمة t المحسوبة من عينة عشوائية من الحجم $n = 20$ تقع في الفترة ($-t_{0.05}$, $t_{0.05}$) ما الاستنتاج الذي يمكن وضعه عند اختبار عينة عشوائية من الحجم $n = 20$ لها متوسط حسابي $\bar{x} = 530$ وانحراف معياري $s = 20$ ؟
 الحل .

من جدول توزيع t في ملحق (٤) نجد أن $t_{0.05} = 1.729$ عند درجات حرية $v = 19$.
 وعلى ذلك فإن المنتج يحقق المواصفات القياسية إذا كانت قيمة t المحسوبة من عينة عشوائية من الحجم $n = 20$ مصباحاً تقع في الفترة (-1.729 , 1.729) . ونجد أن :

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{530 - 500}{\frac{20}{\sqrt{20}}} = 6.708.$$

لا تقع في الفترة (-1.729 , 1.729) إذا كانت $\mu > 500$ فإن قيمة t المحسوبة من العينة تكون جيدة وتعني أن الإنتاج أفضل من المطلوب .

إذا كانت لدينا عنتان عشوائتان مأخوذتان من مجتمعين طبيعيين بمتوسطي μ_1, μ_2 وسوف تكون نظرية (٧-٦) مفيدة فقط إذا كانت العنتان مستقلتان وتبايني المجتمعين σ_1^2, σ_2^2 معلومتان أو يمكن تقديرهما من عينات عشوائية من الحجم n_1, n_2 حيث $n_1 \geq 30$, $n_2 \geq 30$.

إذا كانت σ_1^2, σ_2^2 مجهولتان كما يحدث في معظم الحالات ، فإن التوزيع المضبوط للمتغير Z في نظرية (٧-٦) لا يكون معروفاً فيما عدا لو فرضنا أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ والباين يمكن تقديره من الصيغة التالية :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

وعلى ذلك يكون :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

هي قيمة للإحصاء T الذي يُلضع لتوزيع t بدرجات حرية $v = n_1 + n_2 - 2$.
 نظرية (٧-١٠) إذا كان s_1^2, \bar{x}_1 يمثلان المتوسط والتباين على التوالي لعينة عشوائية من
 الحجم n_1 مأخوذة من مجتمع طبيعي بمتوسط μ_1 وتباين مجهول σ_1^2 وإذا كانت s_2^2, \bar{x}_2
 يمثلان المتوسط والتباين على التوالي لعينة عشوائية من الحجم n_2 مأخوذة من مجتمع طبيعي
 بمتوسط μ_2 وتباين σ_2^2 مجهول وإذا كانت n_1 مستقلة عن n_2 و $\sigma_1^2 = \sigma_2^2$ فإن :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

هي قيمة لطير عشوائي T له توزيع t بدرجات حرية $v = n_1 + n_2 - 2$.
 مثال (٧-١٨) يقوم مصنع لإنتاج نوعين من المصابيح الكهربائية A, B. المصابيح من
 النوع A لها متوسط عمر أطول 100 ساعة عن متوسط عمر المصابيح من النوع B.
 التباين لكلا النوعين واحد. يختار شهريا 15 مصباحا من النوع الأول، 10 مصابيح من
 النوع الثاني للاختبار وتحسب قيمة t. تحقق القيمة الموصفات القياسية إذا وقعت
 في الفترة $(-t_{0.01}, t_{0.01})$ فإذا تم سحب في شهر ما عينة عشوائية من 15 مصباحا من
 المصنع A ووجد أن $s_1 = 50, \bar{x}_1 = 520$. أيضا تم سحب عينة عشوائية من المصنع B
 من الحجم $n_2 = 10$ ووجد أن $\bar{x}_2 = 500, s_2 = 40$ هل الإنتاج يحقق الموصفات
 القياسية ؟

الحل . باستخدام جدول توزيع t في ملحق (٤) فإن $t_{0.01} = 2.5$ بدرجات
 حرية $v = 15 + 10 - 2 = 23$. وعلى ذلك الإنتاج يحقق الموصفات القياسية إذا
 كانت قيمة t الخسوبة تقع في الفترة $(-2.5, 2.5)$. التباين المتجمع s_p^2 هو :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(14)(50)^2 + (9)(40)^2}{15 + 10 - 2} = 2147.826.$$

وبأخذ الجذر التربيعي للتباين المتجمع s_p^2 فإن $s_p = 46.3446$. وعلى ذلك :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \\ = \frac{(520 - 500) - (100)}{46.3446 \sqrt{\frac{1}{15} + \frac{1}{10}}} = -4.228.$$

وبما أن t لا تقع في الفترة (2.5 , -2.5) فإن الإنتاج لا يحقق المواصفات القياسية .

في بعض الأحيان بدلا من استخدام طريقة العينات المستقلة فإنه غالبا ما تستخدم طريقة العينات المتزاوجة **paired samples** . ففي تجارب تغذية الحيوان عند مقارنة عليقتين، حيث توضع الحيوانات المتجانسة في أزواج ويشترط أن تكون هذه الأزواج على درجة عالية من التماثل وقد تختلف الأزواج فيما بينها إلا أن أفراد كل زوج تكون متماثلة ويعطى أحد أفراد كل زوج عليقة بينما يعطى الآخر العليقة الأخرى وبالتالي فإن المقارنة بين العليقتين تتم داخل مجموعات متجانسة. في بعض الأحيان يتم ازدواج المشاهدات لفردات العينة نفسها. فمثلا لمعرفة تأثير دواء على ارتفاع ضغط الدم تختار عينة عشوائية من الحجم n من الأشخاص ويتم قياس ضغط الدم الخاص بهم في أول فترة زمنية ثم يعالجون بهذا الدواء وبعد فترة زمنية معينة يتم قياس ضغط الدم لهم مرة أخرى . أزواج المشاهدات سوف تكون $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. الفروق لأزواج المشاهدات سوف تكون $d_i = x_i - y_i, i = 1, 2, \dots, n$. هذه الفروق تمثل قيم لمغير عشوائي D . متوسط مجتمع الفروق μ_D سوف يساوى الصفر إذا كان الدواء ليس له تأثير . وسوف نرمز لمتوسط الفروق في عينتنا بالرمز \bar{d} ، وسوف تختلف \bar{d} من عينة إلى أخرى ولذلك يعتبر \bar{d} قيمة للإحصاء \bar{D} . كما يحسب تباين الفروق s_D^2 والذي يعتبر قيمة للإحصاء S_D^2 لأنه يتغير من عينة إلى أخرى.

نظرية (٧ - ١١) إذا كان d_1, d_2, \dots, d_n تمثل الفروق لعدد n من أزواج المشاهدات وإذا كانت الفروق التي عددها n تمثل عينة عشوائية لها متوسط \bar{d} وتباين s_D^2 مأخوذة من مجتمع الفروق الطبيعي والذي له متوسط μ_D وتباين σ_D^2 فإن :

$$t = \frac{\bar{d} - \mu_D}{\frac{s_D}{\sqrt{n}}}$$

هي قيمة لمطير عشوائي T له توزيع t بدرجات حرية $v = n - 1$.

مثال (٧ - ١٩) إذا كان من المعتقد أن أكل السمك يساعد على زيادة الذكاء . أجريت تجربة على 11 شخصا تم اختيارهم عشوائيا وأجرى لهم أحد اختبارات الذكاء ثم أعطى لهم طعام يحتوي أساسا على السمك وبعد فترة معينة أجرى لهم اختبار الذكاء مرة أخرى فكانت النتائج كالآتي :

قبل أكل السمك	96	109	104	120	120	100	80	111	90	102	103
بعد أكل السمك	97	112	105	105	117	101	89	114	105	105	109

يفرض أن مستوى الذكاء قبل وبعد أكل السمك يتبع توزيعا طبيعيا وإذا اعتبرنا أن السمك له تأثير على مستوى ذكاء الأشخاص إذا لم تقع قيمة t المحسوبة من العينة في الفترة $(t_{0.05}, -t_{0.05})$ بدرجات حرية $v = n - 1 = 9$ ، ما الاستنتاج الذي يمكن الحصول عليه من بيانات العينة السابقة ؟
الحل .

من جدول t في ملحق (٤) فإن $t_{0.05} = 1.812$ بدرجات حرية $v = 10$. وعلى ذلك يعتبر أكل السمك له تأثير على مستوى ذكاء الأشخاص إذا وقعت قيمة t خارج الفترة $(1.812, -1.812)$. من بيانات العينة فإن الفروقات :

$$-1, -3, -1, 15, 3, -1, -9, -3, -15, -3, -6$$

ومنها :

$$n = 11, \Sigma d_i = -24, \Sigma d_i^2 = 606$$

$$\bar{d} = \frac{\Sigma d_i}{n} = \frac{-24}{11} = -2.1818,$$

$$s_d = \sqrt{\frac{1}{n-1} \left[\Sigma d_i^2 - \frac{(\Sigma d_i)^2}{n} \right]}$$

$$= \sqrt{\frac{1}{10} \left[606 - \frac{(-24)^2}{11} \right]} = 7.44067.$$

$$t = \frac{\bar{d}}{s_d / \sqrt{n}} = \frac{2.1818}{7.44067 / \sqrt{11}} = 0.97252.$$

وحيث أن قيمة t المحسوبة تقع في الفترة (1.812 , -1.812) فيمكن القول أن السهمك ليس له تأثير على مستوى الذكاء .

(٧-٧) توزيع مربع كاي Chi - Square Distribution

إذا تكرّر سحب عينات من الحجم n من توزيع طبيعي تباينه σ^2 وإذا تم حساب تباين العينة s^2 لكل عينة فإننا نحصل على قيم للإحصاء S^2 . التوزيع العيني للإحصاء S^2 له تطبيقات قليلة في الإحصاء . اهتمامنا سوف يكون في توزيع المتغير X^2 والتي نحسب قيمته من الصيغة الآتية :

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} .$$

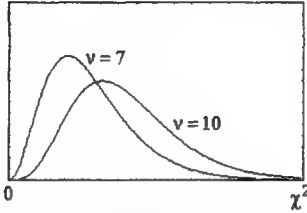
توزيع المتغير العشوائي X^2 يسمى توزيع χ^2 (توزيع مربع كاي) بدرجات حرية $v = n - 1$. كما ذكرنا سابقاً فإن v تساوي المقام في صيغة s^2 .

من الواضح أن قيم χ^2 لا يمكن أن تكون سالبة وعلى ذلك فإن منحنى توزيع χ^2 لا يمكن أن يكون متماثل حول الصفر. التوزيع العيني للإحصاء X^2 يمكن الحصول عليه باختيار عينات عشوائية متكررة من الحجم n من مجتمع طبيعي وحساب القيم χ^2 لكل عينه ثم يمكن الحصول على منحنى χ^2 بتمهيد المدرج التكراري للقيم χ^2 . يعتمد شكل المنحنى على قيم v . يوضح شكل (٧-١٧) منحنيان لتوزيع χ^2 بدرجات حرية $v = 7$ و $v = 10$ حيث يمثل المنحنى بدرجات حرية $v = 7$ توزيع قيم χ^2 المحسوبة من كل العينات من الحجم $n = 8$ من مجتمع طبيعي تباينه σ^2 . بنفس الشكل يمثل المنحنى بدرجات حرية $v = 10$ توزيع كل قيم χ^2 المحسوب من كل العينات من الحجم $n = 11$. نظرية (٧-١٢) إذا كان s^2 يمثل تباين العينة من الحجم n المأخوذة من مجتمع طبيعي له تباين σ^2 فإن :

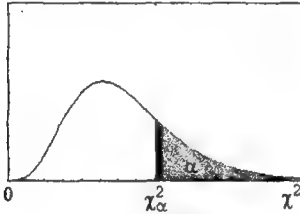
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

هي قيمة لمتغير عشوائي X^2 له توزيع χ^2 بدرجات حرية $v = n - 1$.

شكل (٧-١٧)



بفرض أن χ^2_{α} ترمز لقيمة χ^2 التي توجد على المحور الأفقي تحت منحنى χ^2 بدرجات حرية v والتي تكون المساحة على يمينها قدرها α كما هو موضح في شكل (٧-١٨).



شكل (٧-١٨)

الجدول في ملحق (٥) يعطى قيم χ^2_{α} وذلك لقيم مختلفة من α و v حيث α تأخذ القيم :

.995, .99, .975, .95, .90, .10, .05, 0.025, .01, .005

ودرجات حرية من $v=1$ إلى $v=40$. يوضح الصف الثاني من الجدول قيم χ^2_{α} والعمود الأول من الشمال قيم درجات الحرية أما محتويات الجدول فهي لقيم χ^2_{α} . وعلى ذلك للحصول على قيمة χ^2_{α} بدرجات حرية $v=6$ والتي تكون المساحة على يمينها تساوى .05 فإننا نبحث في الجدول عند تقاطع الصف الذى به $v=6$ مع العمود .05. وعلى ذلك $\chi^2_{.05} = 12.592$. ولعلم تماثل منحنى توزيع χ^2 فلا بد من

استخدام الجدول لإيجاد $\chi^2_{.95} = 1.635$ عند $v = 6$.
 مثال (٧-٢٠) أوجد قيمة χ^2 لتوزيع χ^2 بدرجات حرية $v = 14$ والتي تكون المساحة على يمينها تساوى 0.01 .
 الحل .

بالبحث في جدول توزيع χ^2 في ملحق (٥) عند تقاطع الصف $v = 14$ مع العمود $\alpha = 0.01$ نجد أن $\chi^2 = 29.141$.
 مثال (٧-٢١) أوجد قيمة χ^2 التي تكون المساحة على يسارها تساوى 0.99 لتوزيع χ^2 بدرجات حرية $v = 4$.
 الحل .

قيمة χ^2 التي تكون المساحة على يسارها تساوي 0.99 هي χ والتي تكون المساحة على يمينها تساوى $0.01 = 1 - 0.99$ وعلى ذلك فإن قيمة $\chi^2_{.01}$ بدرجات حرية $v = 4$ هي تلك القيمة في جدول توزيع χ^2 التي تقع عند تقاطع الصف $v = 4$ والعمود $\alpha = 0.01$ هي $\chi^2_{.01} = 13.277$.

مثال (٧-٢٢) أوجد قيمتي χ^2 لتوزيع χ^2 بدرجات حرية $v = 15$ اللتين تحصران بينهما 99% من المساحة تحت المنحنى بحيث أن المساحة في الطرف الأيسر للتوزيع تساوى المساحة في الطرف الأيمن للتوزيع.

الحل . المطلوب هنا هو إيجاد قيمتي χ^2 اللتين تقسمان المنحنى بحيث أن المساحة في الطرف الأيمن هي 0.005 وعلى ذلك قيمة χ^2 في الطرف الأيمن هي $\chi^2_{.005} = 32.799$. وللحصول على قيمة χ^2 في الطرف الأيسر والتي المساحة التي تقع على يسارها هي 0.005 وبالتالي فإن المساحة التي تقع على يمينها هي $0.995 = 1 - 0.005$ وعلى ذلك فإن قيمة χ^2 في الطرف الأيسر هي $\chi^2_{.995} = 4.600$ وبالتالي فإن القيمتين هما 4.600 ، 32.799 .

F Distribution

(٧-٨) توزيع F

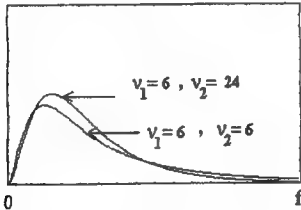
يعتبر توزيع F من التوزيعات الاحتمالية الهامة التي تستخدم في مجال الإحصاء التطبيقي . نظرياً يمكن تعريف توزيع F (توزيع ف) كنسبة لتوزيعين مستقلين يتبعان توزيع χ^2 وكل منهما له درجات حرية خاصة به . لهذا كانت f قيمة للمتغير العشوائي F ، فإن :

وكل منهما له درجات حرية خاصة به . فإذا كانت f قيمة للمتغير العشوائي F ، فإن :

$$f = \frac{\chi_1^2 / v_1}{\chi_2^2 / v_2} = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} = \frac{\sigma_2^2 s_1^2}{\sigma_1^2 s_2^2}.$$

حيث χ_1^2 هي قيمة لتوزيع χ^2 بدرجات حرية $v_1 = n_1 - 1$ و χ_2^2 هي قيمة لتوزيع χ^2 بدرجات حرية $v_2 = n_2 - 1$.

لحساب قيمة f نختار عينة عشوائية من الحجم n_1 من مجتمع طبيعي له تباين σ_1^2 ونحسب s_1^2 / σ_1^2 . أيضا نختار عينة عشوائية مستقلة من الحجم n_2 من مجتمع طبيعي آخر له تباين σ_2^2 ونحسب s_2^2 / σ_2^2 . النسبة للقيمتين s_1^2 / σ_1^2 و s_2^2 / σ_2^2 تنتج قيمة f . توزيع كل القيمة الممكنة من f حيث s_1^2 / σ_1^2 يمثل البسط و s_2^2 / σ_2^2 يمثل المقام يسمى توزيع F بدرجات حرية v_1, v_2 . إذا اعتبرنا كل النسب الممكنة حيث s_2^2 / σ_2^2 يمثل البسط و s_1^2 / σ_1^2 يمثل المقام ، في هذه الحالة نحصل على توزيع كل القيم الممكنة التي تصع توزيع F ولكن بدرجات حرية v_2, v_1 . درجات الحرية المرتبطة بتباين القيممة التي في البسط دائما يوضع أولا متبوعا بدرجات الحرية المرتبطة بتباين العينة التي في المقام . وعلى ذلك منحى توزيع F يعتمد ليس فقط على المعلمتين v_1, v_2 . ولكن أيضا على ترتيبهما وبمجرد الحصول على القيمتين يمكن تعريف المنحنى . منحنيان لتوزيع F موضحة في شكل (١٩-٧) .

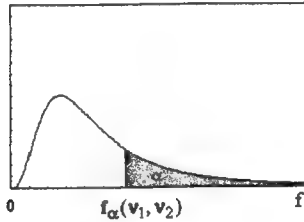


شكل (١٩-٧)

نظرية (١٣-٧) إذا كانت s_1^2 ، s_2^2 تمثلال تباين عينتين عشوائيتين مستقلتين من الحجم n_1 ، n_2 مأخوذتين من مجتمعين من طبيعين بتباين σ_1^2 ، σ_2^2 على التوالي فإن :

$$f = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} = \frac{\sigma_2^2 s_1^2}{\sigma_1^2 s_2^2}.$$

هي قيمة لمغير عشوائي F يتبع توزيع F بدرجات حرية v_1, v_2 بفرض أن $f_\alpha(v_1, v_2)$ ترمز لقيمة f على المحور الأفقي تحت منحنى توزيع F بدرجات حرية $v_1 = n_1 - 1$ و $v_2 = n_2 - 1$ والتي تكون المساحة على يمينها تساوي α والموضحة في شكل (٧-٢٠).



شكل (٧-٢٠)

لاستخراج قيم $f_\alpha(v_1, v_2)$ يوجد جدولان في ملحق (٦) وملحق (٧)، الأول عند $\alpha = 0.05$ والآخر عند $\alpha = 0.01$ وفي كل منهما يكون الصف الأول لقيم v_1 والعمود الأول لقيم v_2 أما محتويات الجدول فهو لقيم $f_\alpha(v_1, v_2)$. على سبيل المثال من جدول توزيع F نلاحظ أن :

$$f_{01}(5,7) = 7.46, \quad f_{05}(1,4) = 7.71$$

$$f_{01}(9,10) = 4.94, \quad f_{05}(4,1) = 224.6$$

وباستخدام النظرية التالية يمكن استخدام جدول توزيع F في إيجاد $f_{1-\alpha}(v_1, v_2)$.

نظرية (٧-١٤) يمكن كتابة $f_{1-\alpha}(v_1, v_2)$ بدرجات حرية v_1, v_2 على الشكل :

$$f_{1-\alpha}(v_1, v_2) = \frac{1}{f_{\alpha}(v_2, v_1)}$$

وعلى ذلك فإن قيمة $f_{.05}(7, 12)$ هي :

$$f_{.95}(7, 12) = \frac{1}{f_{.05}(12, 7)} = \frac{1}{3.57} = 0.2801$$

حيث أن $f_{.05}(12, 7)$ مستخرجة من جدول توزيع F في ملحق (٦) عند مستوى معنوية $\alpha = .05$ ودرجات حرية $v_1 = 12, v_2 = 7$

تمارين :

١- مجتمع محدود يتكون من القيم 2, 4, 6 (أ) أوجد المدرج التكراري للتوزيع

المتوسط للإحصاء \bar{X} عند سحب عينات عشوائية من الحجم $n = 4$ مع الإرجاع (ب)

$$\text{تحقق أن } \mu_{\bar{X}} = \mu \text{ و } \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$$

٢- يتكون مجتمع من القيم 1, 1, 2, 2, 3, 4 (أ) أوجد كل العينات الممكنة من

الحجم $n = 2$ والتي يمكن اختيارها من هذا المجتمع بدون إرجاع (ب) أثبت أن :

$$\mu_{\bar{X}} = \mu \text{ و } \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \frac{N-n}{N-1}$$

٣- إذا كانت 3, 4, 7, 9, 12 تمثل مفردات المجتمع محل الدراسة فإذا سحبت عينة

مكونة من مفردتين :

(أ) أحسب التوقع والتباين للمجتمع .

(ب) حدد العينات الممكنة سحبها بحيث يكون حجم كل عينة مفردتين إذا

كان السحب مع الإرجاع .

(ج) حدد العينات الممكنة سحبها إذا كان السحب بدون إرجاع .

(د) أوجد التوقع والتباين للعينات المسحوبة في (ب) و (ج) .

٤- كم عينة عشوائية مختلفة حجمها $n = 3$ يمكن اختيارها بإحلال ثم بدون إحلال

من مجتمعات محدودة مكونة من $N = 25, N = 20, N = 15$ ثم أوجد احتمال

اختيار كل عينة .

٥- كم عينة عشوائية مختلفة حجمها $n = 2$ يمكن اختيارها بإحلال ثم بدون إحلال من مجتمعات محدودة مكونة من $N = 10$, $N = 15$, $N = 30$ ثم أوجد احتمال اختيار كل عينة .

٦- ما هي قيمة معامل التصحيح للمجتمع المحدود عندما تكون :

$$(أ) \quad n = 10 , N = 100 \quad (ب) \quad n = 5 , N = 200$$

$$(ج) \quad n = 3 , N = 50 \quad (د) \quad n = 15 , N = 300$$

٧- يقوم مصنع بإنتاج مصابيح كهربائية بمتوسط عمر 1900 ساعة وانحراف معياري 200 ساعة. احسب احتمال أن يكون متوسط العينة المبنى على عينة حجمها 100 مصباح أكبر من 1850 وأقل من 1920.

٨- صممت آلة للشرب الرطب بحيث أن كمية الشرب الذي تخرجه يتبع التوزيع الطبيعي بمتوسط $\mu = 8$ أوقية للكوب وانحراف معياري 0.5 أوقية. تختبر الآلة دورياً بأخذ عينة من 9 أكواب وحساب متوسط العينة. إذا كان متوسط كمية الشرب المحسوب من 9 أكواب تقع في الفترة $\bar{x} \pm 2\sigma_{\bar{x}}$ فالآلة تعمل بطريقة صحيحة و غير ذلك لابد من فحص الآلة واتخاذ اللازم . ما هو القرار الذي يجب اتخاذه عند سحب عينة عشوائية من 9 أكواب من الشرب الرطب وكان متوسط كمية الشرب للكوب 8.4 أوقية ؟

٩- إذا كان ضغط الدم لجموعة من الأفراد متوسطة 85 وانحرافه المعياري 5. أخذنا عينة عشوائية حجمها 40 فرداً من هذه المجموعة، أوجد احتمال أن متوسط ضغط الدم في العينة أكبر من 80.

١٠- إذا كان متوسط الدخل الشهري للأسر في إحدى المدن هو 7000 ريال بانحراف معياري 500 ريال. اختبرت عينة عشوائية حجمها 100 أسرة من هذه المدينة أوجد احتمال :

(أ) أن يقل متوسط دخل الأسرة في العينة عن 4800 ريال

(ب) أن يتراوح متوسط دخل الأسرة في العينة بين 6000 , 6500.

-١١- صممت إحدى الشركات سيارة بحيث أن أكبر حولة لها 3000 كم ، وتوسع إلى 30 راكبا . إذا علمت أن أوزان الأشخاص الذين يستعملون هذه السيارة تخضع لتوزيع طبيعي متوسطة 70 كم وانحرافه المعياري 10 كم. أحسب احتمال أن تتحمل أكثر من طاقتها إذا كان مجموع الأشخاص الذين يستعملونها أكبر من 3000 كجم.

-١٢- إذا علمت أن درجات الطلاب في مقرر الإحصاء تخضع لتوزيع طبيعي وسطه 70 وانحرافه المعياري 9. أخذت عينة عشوائية حجمها 15 طالبا. أحسب احتمال أن يزيد متوسط علامات العينة عن 74.

-١٣- إذا كان X متغير عشوائي وسطه 180.5 وانحرافه المعياري 12. من هذا التوزيع أخذت عينة حجمها 36. ما هو احتمال أن يزيد وسط العينة عن 190 ؟

-١٤- إذا كان X متغيرا عشوائيا يمثل أوزان أكياس الطحينه التي تنتجها إحدى المؤسسات وكان X خاضعا لتوزيع طبيعي وسطه 50 كيلو جراما وانحرافه المعياري 5 كيلو جرام. أخذت عينة حجمها 9 أكياس من إنتاج هذه المؤسسة أحسب احتمال أن يزيد الوسط الحسابي لهذه العينة عن 51 كيلو جرام.

-١٥- إذا كان عدد الحوادث الأسبوعية على إحدى الطرق يتبع توزيع بواسون بمتوسط 5 حوادث فإذا أخذت عينة من 60 أسبوعا فما هو احتمال أن يكون متوسط عدد الحوادث فيها أقل من 3 حادث ؟

-١٦- إذا كان متوسط أعمار مرض السكر هو 60 سنة وانحراف معياري 15 سنة. اختبرت عينة عشوائية حجمها 30 مريضا بالسكر. أوجد احتمال :

أ- أن يقل متوسط العمر في العينة عن 50 سنة .

ب - أن يتراوح متوسط العمر في العينة بين 40 , 50 سنة .

-١٧- سحبت عينة عشوائية من الحجم 36 من مجتمع كبير متوسطة $\mu = 80$ وانحرافه المعياري 18 فما هي النسبة المئوية من متوسطات العينات التي :

أ- تقع بين 71 و 89 .

ب- أكبر من 89

ج- تقع بين 74 و 86 د- أقل من 74

١٨- يرغب باحث في تقدير متوسط مجتمع وذلك بالاعتماد على عينة عشوائية مسحوبة من هذا المجتمع وتكون كبيرة بدرجة كافية وذلك تحت شرط :

$$P|\bar{X} - \mu| < \sigma/4 = 0.95$$

ما هي أقل حجم للعينة يجب على الباحث استخدامه ؟

١٩- إذا سحبت كل العينات الممكنة من الحجم 25 من مجتمع طبيعي بمتوسط 50 وانحراف معياري 5. ما هو الاحتمال أن متوسط العينة \bar{X} سوف يقع في الفترة :

$$(\mu_{\bar{X}} - 1.96 \sigma_{\bar{X}}, \mu_{\bar{X}} + 1.96 \sigma_{\bar{X}})$$

٢٠- إذا كانت أطوال 1000 طالب تقريبا تتبع التوزيع الطبيعي بمتوسط 68.5 بوصة وانحراف معياري 2.7 بوصة إذا سحبتنا 200 عينة عشوائية من الحجم $n=25$ من هذا المجتمع وثم حساب المتوسطات . المطلوب :

أ- المتوسط والانحراف المعياري للتوزيع العيني للمتوسط .

ب- عدد متوسطات العينات التي تقع في الفترة (68 , 69.2) .

ج- عدد متوسطات العينات التي تقل عن 67.2 .

٢١- فئاتل حريرية متوسط نقطة قطعها $\mu = 23$ وطلا بانحراف معياري $\sigma = 0.4$ واختيرت عينة عشوائية حجمها 40 فئيلة وذلك لإيجاد نقطة قطعها . ما هو احتمال أن متوسط نقطة القطع للعينة سيكون أكبر من 25.5 ؟

٢٢- في دراسة عن تلوث الهواء بأكسيد الكبريت المتبعث من أحد المصانع كان متوسط الطلوث $\mu = 18$ طنا بانحراف معياري $\sigma = 5.5$ طنا فإذا اختصرت عينة عشوائية مكونة من قراءات 90 يوما أحسب احتمال أن يكون متوسط العينة أكبر من 18.5 .

٢٣- إذا كان X متغيرا عشوائيا يمثل أطوال الطلاب في جامعة ما ، وكان X خاضعا لتوزيع طبيعي بمتوسط 160 سم وانحرافه المعياري 20 سم . أخذت عينة عشوائية حجمها 15 طالبا من هذه الجامعة أحسب احتمال أن يقلل متوسط العينة لأطوال أفراد هذه العينة عن 160 سم .

-٢٤- إذا كان \bar{X} يمثل المتوسط لعينة من الحجم $n_1=2$ مع الإرجاع من مجتمع محدود له القيم 4, 8, 3 بنفس الشكل إذا كان \bar{X}_2 يمثل متوسط عينة من الحجم $n_2=2$ مع الإرجاع من مجتمع 2, 2, 4

أ- أوجد المدرج التكرارى للتوزيع العيني للإحصاء $\bar{X}_1 - \bar{X}_2$

ب - تحقق من أن $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$, $\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

-٢٥- سحبت عينة عشوائية من الحجم $n = 25$ من توزيع طبيعى لسه متوسط 80 وانحراف معيارى 5. أيضا أخذت عينة عشوائية أخرى من الحجم 36 من مجتمع آخر طبيعى له متوسط 75 وانحراف معيارى 3. أوجد الاحتمال أن متوسط العينة المسحوبة من 25 مفردة يزيد عن متوسط العينة المسحوبة من 36 مفردة على الأقل 3.4 .

-٢٦- عينات عشوائية من الحجم 100 سحبت بدون إرجاع من مجتمعين A , B فإذا كان لدينا المعلومات التالية :

$$\mu_1 = 10 , \sigma_1 = 2 , \mu_2 = 8 , \sigma_2 = 1$$

أوجد أ- $E(\bar{X}_1 - \bar{X}_2)$ ب- $\sigma_{\bar{X}_1 - \bar{X}_2}^2$

ج- الاحتمال أن الفرق $\bar{X}_1 - \bar{X}_2$ أقل من 1.5 .

د - الاحتمال أن الفرق $\bar{X}_1 - \bar{X}_2$ أكبر من 1.75 ولكن أقل من 2.5

-٢٧- أجرى بحث على أربعة طلبة لمعرفة هل يدخنون أم لا . فإذا كانت نتائج البحث كالتالى : No, yes, No, No

أ- أوجد نسبة المدخنين في المجتمع .

ب - أوجد التوزيع العيني للإحصاء \hat{P} (نسبة المدخنين) وذلك في حالة السحب بإرجاع أو بدون إرجاع ثم أوجد متوسط وتباين التوزيع إذا كان حجم العينة النى تم اختيارها $n = 2$.

-٢٨- أجرى بحث على 5 أفراد لمعرفة حامل ميكروب معين فكانت النتائج كما يلي حامل للميكروب ، غير حامل ، حامل للميكروب ، غير حامل ، غير حامل المطلوب : أ- نسبة الحاملين للميكروب في المجتمع .

ب - إذا اختيرت نسبة عشوائية من فردين أوجد التوزيع العيني للإحصاء \hat{P} في

حالة السحب بإرجاع والسحب بدون إرجاع .

-٢٩- أجرى بحث لاستطلاع آراء 8 أفراد على منتج ما فكانت النتائج :

yes , yes , yes , yes , yes , No, No, No المطلوب :

أ- نسبة الأفراد المفضلون للمنتج .

ب - إذا اختبرت عينة عشوائية من فردين أوجد التوزيع العيني للإحصاء \hat{P} في

حالة السحب بإرجاع والسحب بدون إرجاع .

-٣٠- إذا كانت نسبة المدخنين من الأفراد الذكور البالغين في إحدى المدن 20%

فإذا اختبرت عينة عشوائية من 200 شخص أوجد احتمال أن يقل عدد المدخنين بينهم عن 3 أشخاص .

-٣١- في شركة كبيرة من 4000 موظف وجد أن الغياب في يوم الاثنين يمثل 8%

لعدة سنوات. إذا اختبرت عينة عشوائية من 1500 موظف في يوم الاثنين. أوجد

احتمال أن تكون نسبة الغياب أكبر من 0.07 .

-٣٢- إذا كان 45% من الأفراد من أصحاب الأملاك في بلد كبير يملكون سيارتين

على الأقل. ما هو الاحتمال في عينة عشوائية من 150 شخص من هذا البلد تكون

نسبة امتلاك سيارتين على الأقل 0.55 .

-٣٣- في مجتمع معين كانت نسبة المعاناة من مرض ما هو 0.2. أخذت عينة

عشوائية من هذا المجتمع حجمها 250 (السحب بإرجاع) أوجد :

أ- احتمال أن تكون النسبة في هذه العينة أقل من 0.3 .

ب - بفرض أن السحب بدون إرجاع وحجم المجتمع 500 أوجد احتمال أن تكون

النسبة في العينة أكبر من 0.4 .

-٣٤- في دولة ما معروف أن 70% من المرشحين في الانتخابات يناهضون قضية ما.

فإذا اختبرت عينة عشوائية من هذه الدولة مكون من 100 مرشح أوجد احتمال أن

تكون نسبة المرشحين الذين يناهضون القضية في العينة بين 80% و 60% .

-٣٥- إذا علم أن نسبة البيض التالف التي تنتجها أحد مراكز إنتاج البيض هي 0.2 .

أشترى شخص 300 بيضة من إنتاج هذا المركز أوجد احتمال أن يجد من بينها 10 بيضة

على الأقل تالفة .

-٣٦- إذا كان لدينا آتين وكانت نسبة العيب للآلة الأولى هو 15% والعيب للآلة الثانية هو 12% . سحبت عينتان من إنتاج الآتين حجمهما 600, 650 وحدة على التوالي . فإذا كانت نسبة العيب للعيبة الأولى هو \hat{P}_1 ونسبة العيب في العينة الثانية \hat{P}_2 . احسب قيمة الاحتمال $P(\hat{P}_1 - \hat{P}_2 \leq 0.2)$.

-٣٧- أوجد : أ - $t_{0.99}$, $v = 10$ ب - $t_{0.05}$, $v = 15$

ج - $t_{0.25}$, $v = 17$ د - $t_{0.99}$, $v = 18$

-٣٨- أوجد قيمة t_{α} عندما $v = 23$ بحيث أن $P(-t_{\alpha} < T < t_{\alpha}) = 0.99$.

-٣٩- أوجد قيمة α التي تحقق الآتي : $t_{\alpha} = 1.33$ و $v = 18$.

-٤٠- أوجد الاحتمال أن المتغير العشوائي T والذي يتبع توزيع t :

أ - أكبر من 1.74 عند $v = 17$ ب - أقل من 1.323 عند $v = 21$

-٤١- يقوم مصنع بإنتاج مصابيح كهربائية متوسط عمرها 500 ساعة . للمحافظة على هذا الإنتاج يقوم المصنع شهريا باختيار 25 مصباحا . إذا وقعت قيمة t الغسوبة في الفترة $(-t_{0.05}, t_{0.05})$ فإن الإنتاج يكون محققا للمواصفات القياسية . ما هو الاستنتاج الذي يمكن اتخاذه عند اختيار عينة لها متوسط $\bar{x} = 518$ وانحراف معياري $s = 40$ ساعة ؟ وذلك تحت فرض أن العينة تم اختيارها من مجتمع طبيعي .

-٤٢- أوجد التين العاشر والتين التسعين P_{90} لتوزيع t بدرجات حرية $v = 15$.

-٤٣- أوجد الربع الأول Q_1 والربع الثالث Q_3 والوسط Q_2 لتوزيع t بدرجات حرية $v = 20$.

-٤٤- إذا كانت الأجور اليومية لعمال إحدى الشركات في مدينة ما تخضع للتوزيع الطبيعي بمتوسط $\mu_1 = 5$ وانحراف معياري $\sigma_1 = 0.5$. وكانت الأجور اليومية لعمال شركة مماثلة في مدينة أخرى تخضع للتوزيع الطبيعي بمتوسط $\mu_2 = 4$ وانحراف معياري $\sigma_2 = 10$. وبفرض أننا سحبنا عينة عشوائية من المجتمع الأول من الحجم $n_1 = 25$ وعينة أخرى من المجتمع الثاني من الحجم $n_2 = 20$ أوجد $P(\bar{X}_1 - \bar{X}_2 \leq 0.75)$.

٤٥- اختبرت عينة عشوائية من مجتمع طبيعي له متوسط $\bar{x}_1 = 20$ وانحراف معياري $s_1 = 2$. كما اختبرت عينة أخرى عشوائية من الحجم $n_2 = 12$ من مجتمع آخر له متوسط $\bar{x}_2 = 24$ وانحراف معياري $s_2 = 6$ فإذا كانت $\mu_2 = 19$ ، $\mu_1 = 22$ و σ_1^2 ، σ_2^2 مجهولتان ولكن تقريباً متساويتان أوجد قيمة t .

٤٦- طبق اختبار للقدرة على التفكير الناقد على مجموعة من المراهقين قبل حضورهم برنامج أعد لهذا الغرض مدته 40 أسبوعاً وبعد حضورهم للبرنامج. فإذا اختبرت عينة عشوائية من الحجم $n = 10$ وكان متوسط الفروق $\bar{d} = 5$ بانحراف معياري $s_{\bar{d}} = 0.5$ أوجد قيمة t .

٤٧- أوجد النقاط التالية من جدولاً توزيع χ^2 في ملحق (٥) مع التوضيح بالرسم :

$$\text{أ- } \chi_{.95}^2, \quad \text{ب- } \chi_{.05}^2 \quad \text{و } v = 1$$

٤٨- لتوزيع χ^2 أوجد

$$\text{أ- } \chi_{.001}^2, \quad \text{ب- } \chi_{.975}^2 \quad \text{و } v = 18, \quad \text{و } v = 29,$$

$$\text{ج- } \chi_{\alpha}^2 \text{ بحث أن } P(X^2 \leq \chi_{\alpha}^2) = .99, \quad v = 4,$$

٤٩- أوجد النقاط التالية من جدول توزيع F في ملحق (٦) مع التوضيح بالرسم

$$f_{.05}(12, 7), \quad f_{.95}(6, 10).$$

٥٠- أوجد الاحتمال أن المتغير العشوائي F أكبر من 2.85 إذا كانت F له توزيع

$$v_1 = 8, v_2 = 12 \text{ بدرجات حرية}$$

٥١- أوجد القيمة $f_{\alpha}(12, 8)$ إذا كانت $P(F < f_{\alpha}(12, 8)) = 0.05$

٥٢- أوجد القيمة α لكل من القيم التالية مع التوضيح بالرسم

$$f_{\alpha}(3, 20) = 3.1, \quad f_{\alpha}(9, 8) = 3.39$$

٥٣- أوجد الاحتمال أنه من عينة عشوائية من 25 مفردة من مجتمع طبيعي له تباين

$$\sigma^2 = 6 \text{ سوف يكون التباين } S^2$$

$$\text{أ- أكبر من } 9.1 \quad \text{ب- بين } 4.62, \quad 10.745$$

-٥٤- إذا كانت s_1^2, s_2^2 تثلان تبايني عنتين عشوائيتين من الحجم $n_1, n_2 = 31$ ،
 $25 =$ تم اختيارهما من مجتمعين طبيعيين . بحيث $\sigma_1^2 = 10, \sigma_2^2 = 15$ أوجد :
 $P(S_1^2 / S_2^2) > 1.26$.

-٥٥- إذا كانت s_1^2, s_2^2 تثلان تبايني عنتين عشوائيتين من الحجم $n_1 = 8, n_2 = 12$ ،
 مسحوبتين من مجتمعين طبيعيين تحت فرض أن $\sigma_1^2 = \sigma_2^2$. أوجد الاحتمال
 $P(S_1^2 / S_2^2 < 4.89)$.

الفصل الثامن

فترات الثقة

Confidence Intervals

(١-٨) مقدمة Introduction

يعبر الاستدلال الإحصائي **statistical inference** فرع في علم الإحصاء يهتم بطرق الاستدلال أو التصميم بشأن المجتمع وذلك بالاعتماد على معلومات يتم الحصول عليها من عينات مختارة من المجتمع . سوف نتناول في هذا الفصل الاستدلال عن معالم مجتمعات مجهولة مثل المتوسط ، النسبة ، الانحراف المعياري ، وذلك بحساب إحصاءات من عينات عشوائية وتطبيق نظرية المعاينة التي ناقشناها في الفصل السابع .

ينقسم فرع الاستدلال الإحصائي إلى فرعين أساسيين : التقدير **estimation** واختبارات الفروض **tests of hypotheses** . سوف تقتصر دراستنا في هذا الفصل على موضوع التقدير بينما موضوع اختبارات الفروض سوف نتناوله في الفصل التاسع . الأمثلة التالية توضح الفرق بين الفرعين . يقوم مصنع بإنتاج قضباناً حديدية ، فإذا اختيرت عينة عشوائية مكونة من 200 قضيب من إنتاج هذا المصنع وقيست أطوالها وتم حساب متوسط طول القضيب في العينة . هذا المتوسط يمكن أن يستخدم لتقدير المعلمة الحقيقية للمجتمع μ . المعلومات عن توزيع المعاينة للإحصاء \bar{X} سوف يساعدنا في حساب درجة الثقة في تقديرنا . هذه المشكلة تنتمي إلى فرع التقدير . الآن إذا كان معروفاً أن جسم الإنسان البالغ يحتاج يومياً في المتوسط إلى 800 ملليجرامات من الكالسيوم لكي يقوم بوظائفه خير قيام . يعتقد علماء التغذية أن الأفراد ذوى الدخل المنخفض لا يستطيعون تحقيق هذا المتوسط . لاختبار ذلك اختيرت عينة عشوائية من 50 شخصاً بالغا من بين ذوى الدخل المنخفض وتم حساب متوسط ما يتناولونه من الكالسيوم يومياً . في هذا المثال لم نحاول تقدير معلمة ولكن بدلاً من ذلك نحاول الوصول إلى قرار صحيح عن الفرض الذي وضعه علماء التغذية . مرة أخرى نعلم على نظرية المعاينة لتمدنا بمقياس للرجحة الثقة في القرار الذي نتخذه .

يتم تقدير معلمة المجتمع إما كتقدير بنقطة **point estimate** أو كتقدير بفترة **interval estimate** . تقدير النقطة لمعلمة مجتمع ما θ هي قيمة وحيدة (مفردة) $\hat{\theta}$ للإحصاء $\hat{\theta}$. على سبيل المثال القيمة \bar{x} للإحصاء \bar{X} ، والمجموعة من عينة عشوائية من الحجم n ، هي تقدير بنقطة لمعلمة المجتمع μ . بنفس الشكل ، $\hat{p} = \frac{x}{n}$ هي تقدير بنقطة للمعلمة الحقيقية p والتي تمثل نسبة صفة ما في مجتمع .

الإحصاء المستخدم لإيجاد تقدير النقطة يسمى المقدر estimator أو دالة القرار decision function. فعلى سبيل المثال دالة القرار S ، والتي تكون دالة في العينة العشوائية، هي مقدر للمعلمة σ . عينات مختلفة تؤدي إلى تقديرات مختلفة.

بفرض أن $\hat{\theta}$ مقدر حيث القيمة $\hat{\theta}$ هي تقدير بنقطة لمعلمة مجتمع مجهولة θ . من المؤكد أننا نرغب في إيجاد التوزيع العيني للإحصاء $\hat{\theta}$ والذي متوسطة يساوي المعلمة التي نرغب في تقديرها. أي مقدر يحقق هذه الخاصية يسمى مقدر غير متحيز unbiased estimator.

تعريف: يقال للإحصاء $\hat{\theta}$ أنه مقدر غير متحيز للمعلمة θ إذا كان $E(\hat{\theta}) = \theta$. إذا كان $\hat{\theta}_1, \hat{\theta}_2$ مقدران غير متحيزان لمعلمة مجتمع θ فإننا نختار المقدر الذي توزيعه الاحتمالي له أقل تباين. وعلى ذلك إذا كان $\sigma_{\hat{\theta}_1}^2 < \sigma_{\hat{\theta}_2}^2$ ، فإننا نقول أن $\hat{\theta}_1$ مقدرًا أكثر كفاءة من $\hat{\theta}_2$.

تعريف: اعتبر كل المقدرات الغير متحيزة لمعلمة θ . يسمى المقدر الذي له أقل تباين بالمقدر الأكثر كفاءة more efficient للمعلمة θ . للتوزيع الطبيعي تم إثبات أن كلا من \bar{X}, \bar{X} (الوسيط) مقدران غير متحيزين لمعلمة المجتمع μ ولكن تباين \bar{X} أقل من تباين \bar{X} . وعلى ذلك كلا التقديرين \bar{X}, \bar{X} (وسيط العينة)، في المتوسط، سوف يكون لهما نفس متوسط المجتمع μ ، ولكن \bar{X} سوف يكون أقرب للمعلمة μ من \bar{X} .

أي تقدير بفترة لمعلمة θ هو فترة على الشكل $a < \theta < b$ حيث a, b تعتمدان على التقدير بنقطة $\hat{\theta}$ لعينة عشوائية خاصة مختارة من المجتمع موضع الدراسة وأيضاً على التوزيع العيني للإحصاء $\hat{\theta}$. على سبيل المثال إذا اختيرت عينة عشوائية تمثل درجات التحصيل في امتحان القبول لخمسين طالبا من المتقدمين للالتحاق في كلية ما وتم الحصول على الفترة 500 و 550 والتي نتوقع أن المتوسط الحقيقي للدرجات التحصيل داخلها. القيمان النهائيان 500 و 550 سوف تعتمدان على متوسط العينة الخسوية \bar{x} وأيضاً على التوزيع العيني \bar{X} . كلما زادت حجم العينة، فإن $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ سوف تقل، كما عرفنا في الفصل السابع، وبالتالي فإن تقديرنا سوف يقترب من المعلمة μ ويؤدي إلى فترة قصيرة.

عينات مختلفة تؤدي إلى قيم مختلفة لـ $\hat{\theta}$ وبالتالي إلى تقديرات بفترة لمعلمة المجتمع θ . بعض هذه الفترات سوف تحتوي على θ والبعض الآخر لا يحتوي على θ . التوزيع العيني للإحصاء $\hat{\theta}$ سوف يساعدنا في إيجاد a, b لكل العينات الممكنة بحيث أن أي نسبة خاصة من هذه الفترات سوف تحتوي على θ . فعلى سبيل المثال، يتم حساب a, b بحيث تكون 0.95 من كل الفترات الممكنة، مع تكرار المعاينة، سوف تحتوي على θ . وعلى ذلك يكون لدينا

احتمال 0.95 لاختيار واحدة من هذه العينات والتي تؤدي إلى فترة تحتوي على θ . هذه الفترة المحسوبة من عينة عشوائية ، تسمى 95% فترة ثقة confidence interval . بمعنى آخر يكون لدينا 95% ثقة أن فترةنا المحسوبة تحتوي على المعلمة θ . عموما توزيع $\hat{\theta}$ سوف يساعدنا في حساب b , a بحيث يكون لأي نسبة خاصة $1 - \alpha$, $0 < \alpha < 1$, من الفترات المحسوبة من كل العينات الممكنة سوف تحتوي على المعلمة θ . الفترة المحسوبة تسمى 100%(1 - α) فترة ثقة للمعلمة θ . تعتبر فترة الثقة الأطول ، هي الأكثر ثقة في الحصول على فترة تحتوي على المعلمة المجهولة . بالطبع يكون من الأفضل الحصول على 95% فترة ثقة أن متوسط العمر لنوع معين من البطاريات ينحصر بين 8 و 5 أسابيع عن الحصول على 99% فترة ثقة أن متوسط العمر ينحصر بين 11 و 2 أسبوعا . دائما يفضل الحصول على فترة قصيرة بدرجة عالية من الثقة . في البنود التالية من هذا الفصل سوف نتناول المقدرات الأكثر كفاءة لعالم المجتمع الشائعة الاستخدام . باستخدام هذه المقدرات وتوزيعاتها الاحتمالية ، يمكن اشتقاق فترات الثقة المقابلة لها .

(٧-٨) فترة ثقة لمتوسط المجتمع μ

Confidence Interval for Population Mean μ

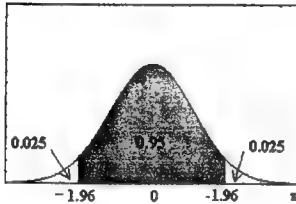
يعتبر الإحصاء \bar{X} هو المقدر الأكثر كفاءة لمعلمة المجتمع μ . التوزيع العيني للإحصاء \bar{X} يتركز عند μ وتباينه أقل من أي مقدر آخر . وعلى ذلك فمتوسط العينة \bar{X} سوف يستخدم كتقدير بنقطة لمتوسط المجتمع μ . مرة أخرى فإن $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ وعلى ذلك فإن العينة الكبيرة سوف تؤدي إلى قيمة للإحصاء \bar{X} ناتجة من توزيع عيني له تباين صغير . وعلى ذلك فإن \bar{X} تقرب من μ عندما تكون n كبيرة .

الآن سوف نتناول الخطوات المتبقية في إيجاد فترة ثقة لمتوسط المجتمع μ وذلك تحت فرض أن العينة مختارة من مجتمع طبيعي أو ، عند عدم تحقق هذا الفرض ، إذا كانت n كبيرة بدرجة كافية . وتبعا للنظريتين (٧-٣) و (٧-٥) فإننا نتوقع أن التوزيع العيني للإحصاء \bar{X} تقريبا يتبع التوزيع الطبيعي بمتوسط $\mu = \bar{X}$ وتباين $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$. تبعا لذلك فإن المتغير العشوائي الطبيعي القياسي Z سوف ينحصر بين 1.96 , 1.96 وذلك باحتمال 0.95 حيث :

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

يتضح من شكل (١-٨) أن :

$$P(-1.96 < Z < 1.96) = 0.95.$$



شكل (١-٨)

وبالتعويض عن Z بقيمتها وبضرب كل حد في المتباينة في $\frac{\sigma}{\sqrt{n}}$ وطرح \bar{X} من كل حد والضرب في -1 نحصل على :

$$P(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}) = 0.95 .$$

وعلى ذلك يمكن القول أنه باحتمال 95% تحتوي الفترة العشوائية $\bar{X} \pm \frac{1.96\sigma}{\sqrt{n}}$ على المعلمة μ . وعلى ذلك نختار عينة عشوائية من الحجم n من المجتمع الذي تباينه σ^2 معلوم ونحسب متوسط العينة \bar{x} وذلك للحصول على 95% فترة ثقة على الشكل :

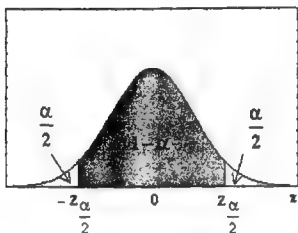
$$\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}} .$$

حيث 1.96 تمثل القيمة الحرجة المقابلة لمعامل الثقة 0.95 وحدتي الثقة هما $\bar{x} + 1.96\sigma/\sqrt{n}$, $\bar{x} - 1.96\sigma/\sqrt{n}$ حيث σ معلومة.

الآن سوف نشرح بصورة عامة طريقة الحصول على $100\%(1-\alpha)$ فترة ثقة حيث $0 < \alpha < 1$. يتضح من شكل (٢-٨) أن $z_{\frac{\alpha}{2}}$ تمثل قيمة z التي تكون المساحة على يمينها $\frac{\alpha}{2}$

تساوي $\frac{\alpha}{2}$. وعلى ذلك يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1 - \alpha .$$



شكل (٢-٨)

وبالتعويض عن Z بقيمتها وبضرب كل حد في المتباينة في $\frac{\sigma}{\sqrt{n}}$ وطرح \bar{X} من كل حد والضرب في -1 نحصل على :

$$P(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha .$$

وعلى ذلك نختار عينة عشوائية من الحجم n من المجتمع الذي تباينه σ^2 معلوم ونحسب متوسط العينة \bar{X} وذلك للحصول على $100\%(1-\alpha)$ فترة ثقة على الشكل :

$$\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} .$$

للعينات العشوائية الصغيرة المختارة من مجتمعات غير طبيعية ، لا نتوقع أن درجة ثقتنا تكون مضبوطة. للعينات من الحجم $n \geq 30$ وبصرف النظر عن شكل المجتمع فإن نظرية المعاينة تؤمن لنا نتائج جيدة.

لحساب $100\%(1-\alpha)$ فترة ثقة للمعلمة μ نفترض أن σ معلومة ولكن عموماً لا يتوافر هذا الفرض ، في هذه الحالة يمكن الاستعاضة عن σ بالانحراف المعياري للعينة s بشرط أن $n \geq 30$.

مثال (٨-١) اختيرت عينة عشوائية من 100 سيجارة من نوع معين وكان متوسط النيكوتين فيها 25 ملليجراماً والانحراف المعياري هو 6 ملليجرامات . أوجد فترة ثقة لمتوسط النيكوتين في السجائر من هذا النوع بدرجة ثقة 95% و 99% .

الحل . التقدير بنقطة للمعلمة μ هو $\bar{x} = 25$ وحيث أن حجم العينة كبير ، فإن الانحراف المعياري للمجتمع σ يمكن الاستعاضة عنه بالانحراف المعياري للعينة $s = 6$. من جدول التوزيع الطبيعي القياسي في ملحق (٣) لأن قيمة z التي على يمينها مساحة قدرها 0.025 وعلى يسارها مساحة قدرها 0.975 هي $z_{0.025} = 1.96$. وعلى ذلك فإن 95% فترة ثقة سوف تكون على الشكل :

$$25 - \frac{(1.96)(6)}{\sqrt{100}} < \mu < 25 + \frac{(1.96)(6)}{\sqrt{100}}$$

والتي تختزل إلى :

$$23.824 < \mu < 26.176$$

للحصول على 99% فترة ثقة وباستخدام جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن قيمة z التي على يمينها مساحة قدرها 0.005 وعلى يسارها مساحة قدرها 0.995 هي $z_{0.005} = 2.575$ وعلى ذلك فإن 99% فترة ثقة تكون على الشكل :

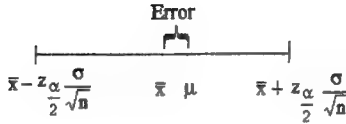
$$25 - \frac{(2.575)(6)}{\sqrt{100}} < \mu < 25 + \frac{(2.575)(6)}{\sqrt{100}}$$

والتي تختزل إلى :

$$23.455 < \mu < 26.545 .$$

تمدنا $100\%(1-\alpha)$ فترة الثقة بتقدير لدقة التقدير بنقطة الذي نحصل عليه . إذا وقعت μ عند مركز الفترة فإن \bar{x} تقدر μ بدون أخطاء. في معظم الحالات فإن \bar{x} لا تساوي μ وبالتالي نحصل على تقدير بنقطة ، \bar{x} ، بخطأ error . حجم هذا الخطأ يساوي الفرق بين \bar{x} و μ . يصل هذا الخطأ إلى القصاة عندما تكون μ قريبة من إحدى حدي الثقة. أي أن \bar{x}

سوف يختلف عن μ بمقدار أقل من $z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ كما يتضح من شكل (٨-٣) .



شكل (٣-٨)

نظرية (١-٨) إذا استخدمت \bar{x} كتقدير لمعلمة المجتمع μ فإنه يكون لدينا $100(1-\alpha)\%$

ثقة أن الخطأ سوف يكون أقل من $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$.

في المثال (١-٨) يكون لدينا 99% ثقة أن متوسط العينة $\bar{x} = 25$ يختلف عن المتوسط الحقيقي μ بمقدار أقل من 1.545.

عادة نرغب في معرفة حجم العينة اللازم للتأكد من أن الخطأ في تقدير μ سوف يكون

أقل من قيمة e . تعني نظرية (١-٨) أننا لا بد من اختيار n بحيث $e = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$.

نظرية (٢-٨) إذا استخدمت \bar{x} كتقدير للمعلمة μ فإن يكون لدينا $100(1-\alpha)\%$ ثقة أن الخطأ سوف يكون أقل من قيمة معينة e وذلك عندما يحسب حجم العينة من الصيغة التالية :

$$n = \left(\frac{z_{\alpha/2} \sigma}{e} \right)^2.$$

الصيغة السابقة تمكن المرء في تحديد مدي كبر العينة التي يحتاج إليها كي يقدر μ لأي درجة يرغبها من درجات الدقة قبل أخذ أي عينة واحدة شريطة أن تكون قيمة σ معلومة . أما إذا لم يكن المرء على علم بقيمة σ ، فلا بد من أخذ عينة مبدئية ، $n \geq 30$ ، كي يحصل على تقدير للمعلمة σ يمكن استخدامه في الصيغة السابقة لتحديد مدي كبر n الواجب .

مثال (٢-٨) ما هو حجم العينة المطلوب في مثال (١-٨) وذلك للحصول على 95% ثقة أن تقديرنا الذي نحصل عليه يختلف عن μ بقيمة أقل من واحد صحيح.

الحل . الانحراف المعياري $s = 6$ والحسوب من عينة عشوائية من الحجم $n = 100$ سوف

تستخدم s بدلا من σ . من نظرية (٢-٨) فإن :

$$n = \left[\frac{(1.96)(6)}{1} \right]^2 = 138.3 \approx 138.$$

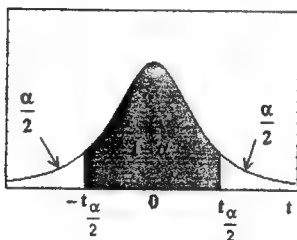
وعلى ذلك يمكن القول أن لدينا 95% ثقة أن العينة العشوائية من الحجم 138 سوف نمثّل بتقدير \bar{x} يختلف عن μ بقيمة أقل من الواحد الصحيح .

في معظم الأحيان يكون المطلوب تقدير متوسط المجتمع عندما يكون التباين غير معلوم وحجم العينة أقل من 30، فقد تكون التكاليف عاملاً محدداً لحجم العينة. طالما كان شكل المجتمع (تقريباً) ناقوسياً فإنه يمكن حساب فترات الثقة عندما تكون σ^2 غير معلومة وحجم العينة صغير وذلك باستخدام التوزيع العيني للمتغير T ، حيث أن :

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}.$$

طريقة إيجاد $100(1 - \alpha)\%$ فترة ثقة في هذه الحالة هي نفسها الطريقة المتبعة في حالة العينات الكبيرة فيما عدا استخدام توزيع t بدلا من التوزيع الطبيعي القياسي. .
يتضح من شكل (٤-٨) أن :

$$P\left(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}\right) = 1 - \alpha.$$



شكل (٤-٨)

حيث أن $t_{\frac{\alpha}{2}}$ هي قيمة t بدرجة حرية $v = (n-1)$ والتي تكون المساحة على يمينها تساوي $\frac{\alpha}{2}$.

ونظراً لخاصية التماثل لتوزيع t فإن مساحة مساوية قدرها $\frac{\alpha}{2}$ تقع على يسار القيمة $-t_{\frac{\alpha}{2}}$.

بالتعويض عن T بقيمتها فإننا يمكن كتابة :

$$P(-t_{\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} < t_{\frac{\alpha}{2}}) = 1 - \alpha.$$

وبضرب كل حد في المتباينة في $\frac{S}{\sqrt{n}}$ وطرح \bar{X} من كل حد والضرب في -1 نحصل على :

$$P(\bar{X} - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}) = 1 - \alpha.$$

لعينة خاصة من الحجم n ، بحسب المتوسط \bar{x} والانحراف المعياري s ويتم الحصول على 100% $(1 - \alpha)$ فترة ثقة كما يأتي :

$$\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}.$$

مثال (٣-٨) في اختبار للزمن الذي يستغرقه تجميع ماكينة معينة وجد أن الزمن الذي استغرقه تجميع 6 ماكينات هو على التوالي : 12, 13, 11, 5, 10, 12 (مقاسه بالدقائق). أوجد 95% فترة ثقة للمتوسط المجتمع μ وذلك تحت فرض أن الزمن (في هذا المثال) يتبع توزيعاً طبيعياً.

الحل . متوسط العينة والانحراف المعياري للبيانات المعطاة هما : $\bar{x} = 10.5$, $s = 2.881$. باستخدام جدول توزيع t في ملحق (٤) فإن $t_{0.025} = 2.571$ وذلك عند درجات حرية $v = 5$. وعلى ذلك 95% فترة ثقة للمعلمة μ هي :

$$10.5 - \frac{(2.571)(2.881)}{\sqrt{6}} < \mu < 10.5 + \frac{(2.571)(2.881)}{\sqrt{6}}$$

التي تختزل إلى :

$$7.476 < \mu < 13.524.$$

(٣-٨) فترة ثقة للفرق بين متوسطي مجتمعين $\mu_1 - \mu_2$

Confidence Interval for the Difference Between two Populations Means

إذا كان لدينا مجتمعان ، المجتمع الأول له متوسط μ_1 وتباين σ_1^2 والمجتمع الثاني له متوسط μ_2 وتباين σ_2^2 . التقدير الأكثر كفاءة للفرق بين متوسطين $\mu_1 - \mu_2$ هو الإحصاء $\bar{X}_1 - \bar{X}_2$. وعلى ذلك ، للحصول على تقدير بنقطة للمعلمة $\mu_1 - \mu_2$ لا بد من اختيار عينة عشوائية من الحجم n_1 من المجتمع الأول وعينة عشوائية من الحجم n_2 من المجتمع الثاني ومستقلة عن العينة الأولى وحساب الفرق بين المتوسطين $\bar{X}_1 - \bar{X}_2$. بفرض أن العيّتين المستقلين تم اختيارهما من مجتمعين طبيعيين ، أو في حالة عدم توافر ذلك الفرض ، إذا كان كلا من n_1 و n_2 أكبر من أو يساوي 30 فإنه يمكن إيجاد فترة ثقة للمعلمة $\mu_1 - \mu_2$ تعتمد على التوزيع العيني للإحصاء $\bar{X}_1 - \bar{X}_2$.

بالرجوع إلى نظرية (٦-٧) فإننا نتوقع أن التوزيع العيني للإحصاء $\bar{X}_1 - \bar{X}_2$ تقريبا يتبع التوزيع الطبيعي بمتوسط :

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$$

والحرفاء معياري :

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

وعلى ذلك فإن المتغير العشوائي الطبيعي القياسي :

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

سوف يقع بين $z_{\frac{\alpha}{2}}$ و $-z_{\frac{\alpha}{2}}$ باحتمال $1 - \alpha$. وبالرجوع مرة أخرى إلى شكل (٢-٨)

فإنه يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1 - \alpha .$$

وباستبدال Z بقيمتها فإن :

$$P \left[-z_{\frac{\alpha}{2}} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < z_{\frac{\alpha}{2}} \right] = 1 - \alpha.$$

ويضرب كل حد في المتباينة في $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ وطرح $\bar{X}_1 - \bar{X}_2$ من كل حد والضرب في 1 - نحصل على :

$$P \left[(\bar{X}_1 - \bar{X}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < \right.$$

$$\left. (\bar{X}_1 - \bar{X}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right] = 1 - \alpha.$$

لأى عيّنتين عشوائيتين مستقلتين من الحجم n_1, n_2 مأخوذتين من مجتمعين تابيهما σ_1^2, σ_2^2 معلومتين ، فإن $\bar{X}_1 - \bar{X}_2$ تحسب ويمكن إيجاد 100% $(1 - \alpha)$ فترة ثقة كالتالي :

$$(\bar{X}_1 - \bar{X}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

درجة الثقة تكون مضبوطة عندما تختار العينات من مجتمعات طبيعية. للمجتمعات الغير طبيعية يمكن الحصول على فترات ثقة تقريبية والتي تكون جيدة جدا عندما n_1, n_2 تزيد عن 30. إذا كانت σ_1^2, σ_2^2 مجهولين والعينات المختارة كبيرة بدرجة كافية ، فإنه يمكن استبدال $\sigma_1^2, \sigma_2^2 \rightarrow s_1^2, s_2^2$ على التوالي بدون التأثير على فترة الثقة .

مثال (٤-٨) أعطى اختبار في مادة الإحصاء إلى 75 طالبة و 50 طالباً. فإذا كان متوسط النقاط من عينة الطالبات $\bar{x}_1 = 80$ بالحدود المعياري $s_1 = 7$. وكان متوسط النقاط لعينة الطلبة $\bar{x}_2 = 70$ بالحدود المعياري $s_2 = 6$. أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$.

الحل . التقدير بنقطة لـ $\mu_1 - \mu_2$ هو $\bar{x}_1 - \bar{x}_2 = 80 - 70 = 10$. وحيث أن كلا من n_1, n_2 كبيرة فإنه يمكن استخدام $s_1 = 7$ بدلا من σ_1 و $s_2 = 6$ بدلا من σ_2 . باستخدام $\alpha = 0.05$ فإن $z_{0.025} = 1.96$ وذلك من جدول التوزيع الطبيعي في ملحق (٣). وبالتعويض في 95% فترة ثقة لـ $\mu_1 - \mu_2$ التالية :

$$(\bar{x}_1 - \bar{x}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2$$

$$< (\bar{x}_1 - \bar{x}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

نحصل على 95% فترة ثقة على الشكل :

$$10 - 1.96 \sqrt{\frac{7^2}{75} + \frac{6^2}{50}} < \mu_1 - \mu_2 < 10 + 1.96 \sqrt{\frac{7^2}{75} + \frac{6^2}{50}}$$

أو

$$7.703 < \mu_1 - \mu_2 < 12.297.$$

تستخدم الطريقة السابقة لتقدير الفرق بين متوسطين إذا كان σ_1^2, σ_2^2 معلومتان أو يمكن تقديرهما من عينات كبيرة. إذا كانت أحجام العينات صغيرة ، لابد من استخدام توزيع t للحصول على فترات ثقة والتي تكون صحيحة عندما تكون المجتمعات تقريبا تتبع التوزيع الطبيعي.

فرض أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ فإنه يمكن استخدام s_p^2 كتقدير للتباين العام σ^2 . وباستخدام نظرية (٧-١٠) ، يتضح من شكل (٤-٨) أن :

$$P(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}) = 1 - \alpha.$$

حيث :

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\left(\frac{1}{n_1}\right) + \left(\frac{1}{n_2}\right)}}$$

$t_{\frac{\alpha}{2}}$ هي قيمة t ، بدرجات حرية $v = n_1 + n_2 - 2$ ، والتي المساحة على يمينها تساوي $\frac{\alpha}{2}$.

باستبدال T بقيمة يمكن كتابة :

$$P \left[(\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 \right. \\ \left. < (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right] = 1 - \alpha.$$

لأي عيتين عشوائيتين مستقلتين من الحجم n_1 ، n_2 يتم اختيارهما من مجتمعين طبيعيين فإن الفرق بين متوسطي العيتين، $\bar{X}_1 - \bar{X}_2$ ، والتباين العام للعينة s_p^2 يتم حسابهما واستخدامهما في إيجاد $100(1 - \alpha)\%$ فترة ثقة لـ $\mu_1 - \mu_2$ على الشكل :

$$(\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 \\ < (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

مثال (٨-٥) اختبرت مجموعتان من الأراب ، الأولى من 13 أرنباً وأعطيت الغذاء A والثانية من 15 أرنباً وأعطيت الغذاء B وكانت الزيادة في الوزن بعد فترة معينة هي :

A: 35, 30, 30, 23, 21, 12, 24, 23, 33, 27, 29, 25, 21.

B: 20, 17, 34, 31, 29, 39, 30, 46, 7, 21, 33, 43, 21, 34, 20.

أوجد 95% فترة ثقة للفرق بين متوسطي المجتمعين ، وذلك تحت فرض أن المجتمعين تقريباً

يتبعان التوزيع الطبيعي حيث $\sigma_1^2 = \sigma_2^2$.

الحل .

$$n_1 = 13, \quad \bar{X}_1 = 25.62, \quad s_1 = 6.05,$$

$$n_2 = 15, \quad \bar{X}_2 = 28.33, \quad s_2 = 10.58,$$

لإيجاد فترة ثقة لـ $\mu_1 - \mu_2$ سوف نستخدم التقدير بنقطة $\bar{X}_1 - \bar{X}_2 = 25.62 - 28.33 = -2.71$

التباين العام s_p^2 هو :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(12)(6.05)^2 + (14)(10.58)^2}{13 + 15 - 2} \\ = 77.1669.$$

بأخذ الجذر التربيعي للتباين العام فإن $s_p = 8.784$. باستخدام $\alpha = .025$ فإن $t_{.025} = 2.056$.
تستخرج من جدول توزيع t في ملحق (٤) عند درجات حرية $26 = 13 + 15 - 2$.
بالتعويض في الصيغة التالية :

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2$$

$$< (\bar{x}_1 - \bar{x}_2) + t_{\alpha} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

فإن 95% فترة ثقة لـ $\mu_1 - \mu_2$ هي :

$$-2.71 - (2.056)(8.784) \sqrt{\frac{1}{13} + \frac{1}{15}} < \mu_1 - \mu_2 < \\ -2.71 + (2.056)(8.784) \sqrt{\frac{1}{13} + \frac{1}{15}}.$$

والتي يمكن اختزالها إلى :

$$-9.553 < \mu_1 - \mu_2 < 4.133.$$

تفترض الطريقة السابقة للحصول على فترات ثقة لـ $\mu_1 - \mu_2$ أن المجتمعين طبيعيين وأن

$$\sigma_1^2 = \sigma_2^2. \text{ أيضا يمكن الحصول على نتائج جيدة إذا كانت } \sigma_1^2 \neq \sigma_2^2 \text{ وذلك تحت}$$

شرط أن المجتمعين طبيعيين و $n_1 = n_2$.

الآن وعند الرغبة في إيجاد 100%(1- α) فترة ثقة لـ $\mu_1 - \mu_2$ في حالة العينات

الصغيرة عندما تكون $\sigma_1^2 \neq \sigma_2^2$ وعند صعوبة الحصول على عينات ذات أحجام متساوية.

الإحصاء الأكثر استخداماً في هذه الحالة هو :

$$T' = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{S_1^2}{n_1}\right) + \left(\frac{S_2^2}{n_1}\right)}}$$

والذي تقريباً يتبع توزيع t بدرجات حرية v ، حيث :

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 - 1}}$$

وبما أن v نادراً ما تكون عدد صحيح ، فإننا نقرنها إلى أقرب رقم صحيح. وعلى ذلك يمكن كتابة :

$$P\left(-t_{\frac{\alpha}{2}} < T' < t_{\frac{\alpha}{2}}\right) \approx 1 - \alpha.$$

حيث $t_{\frac{\alpha}{2}}$ هي قيمة لتوزيع t ، بدرجات حرية v ، والتي المساحة على يمينها تساوي $\frac{\alpha}{2}$.

بالتعويض عن T' بقيمتها في المتباينة وإتباع الخطوات نفسها المتبعة في الحالات السابقة يمكن الحصول على $100(1 - \alpha)\%$ فترة ثقة لـ $\mu_1 - \mu_2$ كالآتي :

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2$$

$$< (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

مثال (٨ - ٦) خلال 20 سنة ماضية كان متوسط سقوط المطر في المنطقة A في قطر ما خلال شهر يناير 1.8 بوصة بانحراف معياري 0.4 بوصة. بينما كان متوسط سقوط المطر في المنطقة B من نفس القطر خلال 15 سنة ماضية 1.03 بوصة بانحراف 0.25 بوصة . أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$ وذلك تحت فرض أن المفردات مأخوذة من مجتمعين طبيعيين حيث $\sigma_1^2 \neq \sigma_2^2$.

الحل . المنطقة A : فإن $n_1 = 20$, $\bar{x}_1 = 1.8$, $s_1 = 0.4$,
وللمنطقة B فإن $n_2 = 15$, $\bar{x}_2 = 1.03$, $s_2 = 0.25$,

وعلى ذلك 95% فترة ثقة لـ $\mu_1 - \mu_2$ حيث $\sigma_1^2 \neq \sigma_2^2$ و $n_1 \neq n_2$ يمكن الحصول عليها بالاعتماد على توزيع t بدرجة حرية تحسب كالآتي :

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\left[\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 - 1} \right]}$$

$$= \frac{\left(\frac{.4^2}{20} + \frac{.25^2}{15} \right)^2}{\left[\frac{\left(\frac{.4^2}{20} \right)^2}{19} + \frac{\left(\frac{.25^2}{15} \right)^2}{14} \right]} = 32.11 \approx 32.$$

وعلى ذلك $\bar{x}_1 - \bar{x}_2 = 1.8 - 1.03 = 0.77$ ونحت فرض أن $\alpha = 0.05$ ومن الجدول في ملحق (t) فإن $t_{0.025} = 2.042$ بدرجات حرية $v = 32$. وبالتعويض في الصيغة التالية :

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2$$

$$< (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} .$$

يمكن إيجاد 95% فترة ثقة (تقريبية) كالآتي :

$$0.77 - 2.042 \sqrt{\frac{.4^2}{20} + \frac{.25^2}{15}} < \mu_1 - \mu_2 <$$

$$0.77 + 2.042 \sqrt{\frac{.4^2}{20} + \frac{.25^2}{15}} .$$

والتي نحصل إلى :

$$0.545 < \mu_1 - \mu_2 < 0.995 .$$

وأخيراً سوف نناقش في هذا البند طريقة تقدير الفرق بين متوسطين عندما تكون العينتين غير مستقلتين. فعلى سبيل المثال عندما تأخذ عينة واحدة ونحصل على قراءات المفردات ثم نضع هذه العينة تحت مؤثر ونعود وتأخذ قراءات أخرى لها ، وبمقارنة مجموعتي القراءتين لنفس المفردات يمكننا استنتاج تأثير هذا العامل أو المؤثر. لنفرض مثلاً أننا نريد معرفة تأثير دواء على قراءات ضغط الدم المرتفع وأخذنا لذلك عينة من 10 شخصاً وقراءنا ضغط الدم لكل منهما ثم أعطينا كل شخص دواء له تأثير على ضغط الدم المرتفع وأعدنا أخذ القراءات مرة أخرى. في هذه الحالة نقول أننا أمام عيتين غير مستقلتين أو عيتين مزدوجتين **paired samples**. أزواج المشاهدات سوف تكون $(x_1 - y_1), (x_2 - y_2), \dots, (x_n - y_n)$ الفروق لأزواج المشاهدات سوف تكون $d_1 = (x_1 - y_1), d_2 = (x_2 - y_2), \dots, d_n = (x_n - y_n)$. هذه الفروق تمثل قيم للمتغير العشوائي **D** التقدير بنقطة لتوسط المجتمع $\mu_D = \mu_1 - \mu_2$ يعطى من \bar{d} والذي يساوى متوسط الفروق في العينة. وبما أن \bar{d} تمثل قيمة للإحصاء \bar{D} كما أن التباين للفروق هو s_d^2 حيث :

$$s_d^2 = \frac{1}{n-1} \left[\sum_{i=1}^n d_i^2 - \frac{(\sum_{i=1}^n d_i)^2}{n} \right]$$

يمثل قيمة للمتغير العشوائي S_d^2 فإن $100\%(1-\alpha)$ فترة ثقة للمعلمة μ_D يمكن الحصول عليها باستخدام نظرية (٧-١١) والتي تسمح بكتابة :

$$P(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}) = 1 - \alpha.$$

حيث :

$$T = \frac{\bar{D} - \mu_D}{\frac{S_d}{\sqrt{n}}}$$

و $t_{\frac{\alpha}{2}}$ هي قيمة لتوزيع t ، بدرجات حرية $v = n - 1$ والتي تكون المساحة على يمينها تساوى $\frac{\alpha}{2}$

الآن سوف نضع نفس الخطوات المتبعة في الحالات السابقة وذلك للحصول على $\frac{\alpha}{2}$.
 $100\%(1-\alpha)$ فترة ثقة كالتالي :

$$\bar{d} - t_{\alpha/2} \frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\alpha/2} \frac{s_d}{\sqrt{n}}.$$

حيث \bar{d} و s_d هما المتوسط والانحراف المعياري للفروق لعدد n من أزواج المشاهدات و $t_{\alpha/2}$

هي قيمة لتوزيع t بدرجات حرية $v = n - 1$ والتي تكون المساحة على يمينها تساوي $\frac{\alpha}{2}$.

مثال (٧-٨) أخذت عينة عشوائية من 10 تلاميذ من إحدى المدارس ودونست أوزانهم ثم أعطي كل منهم كوباً من اللبن صباحاً وآخر ظهراً وذلك لمدة ثلاثة شهور متتالية . ثم دونت أوزانهم فكانت النتائج كالآتي :

الوزن قبل تعاطي اللبن	129	124	126	139	133	136	139	135	137	140
الوزن بعد تعاطي اللبن	130	126	129	140	136	134	141	140	138	141

المطلوب إيجاد 99% فترة ثقة للفروق الحقيقي $\mu_D = \mu_1 - \mu_2$.

الحل . التقدير بنقطة لـ μ_D هو $\bar{d} = -1.7$. التباين s_d^2 لفروق العينة هو :

$$s_d^2 = \frac{1}{n-1} \left[\sum d_i^2 - \frac{(\sum d_i)^2}{n} \right]$$

$$= \frac{1}{9} \left[59 - \frac{(-17)^2}{10} \right] = 3.344.$$

وبأخذ الجذر التربيعي للمقدار $s_d = 1.829$ فإن $\alpha = 0.01$ باستخدام $t_{0.005} = 3.25$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $v = n - 1 = 9$. وبالتعويض في الصيغة التالية :

$$\bar{d} - t_{\alpha/2} \frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\alpha/2} \frac{s_d}{\sqrt{n}}.$$

نحصل على 99% فترة ثقة كالآتي :

$$-1.7 - (3.25) \frac{(1.829)}{\sqrt{10}} < \mu_D < -1.7 + (3.25) \frac{(1.829)}{\sqrt{10}}$$

والتي تختزل إلى :

$$-3.58 < \mu_D < 0.18.$$

(٨-٤) فترة ثقة للنسبة Confidence Interval for Proportion

أوضحنا في الفصل السابع أن الإحصاء \bar{X} ليس دائماً المطلوب في مجال الإحصاء. ففي بعض الأحيان يكون الاهتمام بمعرفة نسبة وجود صفة معينة في مجتمع ما مثل نسبة المصابين بتسوس الأسنان أو نسبة النباتات المصابة وهكذا. التقدير بنقطة الأكثر كفاءة لنسبة صفة ما p هو الإحصاء $\hat{P} = \frac{X}{n}$. وعلى ذلك، فإن نسبة العينة $\hat{p} = \frac{x}{n}$ سوف تستخدم كتقدير بنقطة للمعلمة p . من الفصل السابق ومن نظرية (٧-٧) نعلم أن \hat{p} تقريباً تتبع توزيعاً طبيعياً متوسطه $\mu_{\hat{p}} = p$ وتباينه $\sigma_{\hat{p}}^2 = \frac{pq}{n}$. وعلى ذلك فإن :

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1 - \alpha .$$

حيث :

$$Z = \frac{\hat{P} - p}{\sqrt{\frac{pq}{n}}}$$

و $z_{\frac{\alpha}{2}}$ هي قيمة للمتغير العشوائي الطبيعي القياسي Z والتي المساحة على يمينها تساوي $\frac{\alpha}{2}$.

بالعويض عن Z بقيمتها فإنه يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}} < \frac{\hat{P} - p}{\sqrt{\frac{pq}{n}}} < z_{\frac{\alpha}{2}}) = 1 - \alpha .$$

وبضرب كل حد في المتباينة في $\sqrt{\frac{pq}{n}}$ وطرح \hat{p} والضرب في -1 ، نحصل على :

$$P(\hat{P} - z_{\frac{\alpha}{2}} \sqrt{\frac{pq}{n}} < p < \hat{P} + z_{\frac{\alpha}{2}} \sqrt{\frac{pq}{n}}) = 1 - \alpha .$$

وحيث أن p عادة ما تكون مجهولة لذلك نستخدم \hat{p} بدلاً من p وعلى ذلك فإن :

$$P(\hat{P} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{P} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}) = 1 - \alpha .$$

لعينة عشوائية خاصة من الحجم n (بما لقواعد Cochran) كما ذكرنا في الفصل السابع

نحسب نسبة العينة $\hat{p} = \frac{x}{n}$ ويتم الحصول على $100(1-\alpha)\%$ فترة ثقة للمعلمة p كمايلي :

$$\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} .$$

مثال (٨-٨) يقوم مصنع بإنتاج منتج على درجة عالية من الجودة ويرغب المسئول في المصنع في تقدير نسبة الوحدات المنتجة التالفة. فإذا اختبرت عينة عشوائية من 200 وحدة ووجد أن بينهم 40 وحدة تالفة ، أوجد 95% فترة ثقة للمعلمة p .

الحل . التقدير بنقطة للمعلمة p هو $\hat{p} = \frac{40}{200} = 0.2$. باستخدام جدول التوزيع الطبيعي في

ملحق (٣) فإن $z_{0.025} = 1.96$. بالصيغ في الصيغة التالية :

$$\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} .$$

يمكن الحصول على 95% فترة ثقة كالتالي :

$$0.2 - 1.96 \sqrt{\frac{(0.2)(0.8)}{200}} < p < 0.2 + 1.96 \sqrt{\frac{(0.2)(0.8)}{200}}$$

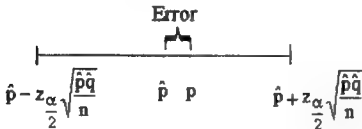
والتي يمكن اختصارها الى :

$$0.145 < p < 0.255 .$$

إذا وقعت p عند مركز $100(1-\alpha)\%$ فترة الثقة ، فإن \hat{p} سوف تقدر p بدون

أخطاء. في معظم الأحوال ، \hat{p} لا تساوي p وعلى ذلك يكون هناك فرق بين \hat{p} و p والذي يمثل الخطأ. هذا الخطأ يصل الى أقصاه عندما تكون p قريبة من إحدى حدي الثقة. وعلى ذلك

\hat{p} سوف تختلف عن p بقيمة أقل من $z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$ كما يتضح من شكل (٨-٥) .



نظرية (٣-٨) إذا استخدمت \hat{p} كتقدير للمعلمة p فإنه يكون لدينا $100\%(1-\alpha)$ ثقة أن الخطأ سوف يكون أقل من $z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$.

عند الرغبة في تقدير حجم العينة اللازم للتأكد من أن الخطأ في تقدير p أقل من مقدار معين e وتبعاً لنظرية (٣-٨) فلا بد من اختيار n بحيث تكون $e = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$.

نظرية (٤-٨) إذا استخدمت \hat{p} كتقدير بنقطة للمعلمة p فإنه يكون لدينا $100\%(1-\alpha)$ ثقة أن الخطأ سوف يكون أقل من قيمة معينة e عندما يحسب حجم العينة من الصيغة التالية :

$$n = \frac{z_{\alpha/2}^2 \hat{p}\hat{q}}{e^2}$$

وحيث أن \hat{p} في الصيغة السابقة لا بد أن نقرر من عينة ، لذلك لا بد من اختيار عينة مبدئية كبيرة وحساب نسبة العينة \hat{p} منها.
مثال (٩-٨) أحسب القيمة العظمى للخطأ في التقدير e عند 95% للمثال (٨-٨).
الحل . حيث e تعطى بالعلاقة :

$$e = z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

وحيث أن $z_{\alpha/2} = 1.96$ ، $\hat{p} = 0.2$ ، $\hat{q} = 0.8$ فإن :

$$e = 1.96 \sqrt{\frac{(0.2)(0.8)}{200}} = 0.0554.$$

مثال (١٠-٨) في عينة عشوائية من 500 مواطن في مجتمع سكاني ما ، وجد منهم 270 مواطناً يحبون أن يضاف الى مياههم قليل من الفلور . المطلوب :

(أ) إيجاد 95% فترة ثقة لنسبة المجتمع الذين يحبون إضافة الفلور .
(ب) تقدير حجم العينة التي يمكننا التأكد منها باحتمال 95% من أن الخطأ لا يتجاوز 0.05.

الحل . (أ) التقدير بنقطة للمعلمة p هو $\hat{p} = \frac{x}{n} = \frac{270}{500} = 0.54$. باستخدام جدول التوزيع

الطبيعي القياسي في ملحق (٣) فإن $z_{0.025} = 1.96$ وبالتعويض في الصيغة التالية :

$$\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

يمكن الحصول على 95% فترة ثقة كالتالي :

$$0.54 - 1.96 \sqrt{\frac{(0.54)(0.46)}{500}} < p < 0.54 + 1.96 \sqrt{\frac{(0.54)(0.46)}{500}}$$

والتي تختزل إلى :

$$0.496 < p < 0.584.$$

(ب) باعتبار الأشخاص الذين عددهم 500 يمثلون عينة عشوائية مبدئية حيث أن $\hat{p}=0.54$
وباستخدام نظرية (٨-٤) فإن :

$$n = \frac{(1.96)^2 (0.54)(0.46)}{(0.05)^2} = 381.70$$

$$\approx 382.$$

(٨-٥) فترة ثقة للفرق بين نسبتي

Confidence Interval for the Difference Between Two Proportions

المقدر بنقطة الأكثر كفاءة للفرق بين نسبتي ، $p_1 - p_2$ ، هو الإحصاء $\hat{p}_1 - \hat{p}_2$. وعلى ذلك ، للحصول على تقدير بنقطة لـ $p_1 - p_2$ سوف تختار عيتين عشوائيتين مستقلتين من الحجم n_1, n_2 وحساب النسبة للصفة موضع الدراسة في كل عينة ، $\hat{p}_1 = \frac{x_1}{n_1}, \hat{p}_2 = \frac{x_2}{n_2}$ ، حيث x_1, x_2 يمثلان عدد المفردات الذين يملكون الصفة موضع الاهتمام في العيتين على التوالي. يتم حساب الفرق $\hat{p}_1 - \hat{p}_2$. فترة ثقة لـ $\hat{p}_1 - \hat{p}_2$ يمكن الحصول عليها بالاعتماد على التوزيع المعيني للإحصاء $\hat{p}_1 - \hat{p}_2$ والذي تقريباً ، تبعاً لنظرية (٧-٨) ، يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{p}_1 - \hat{p}_2} = \mu_1 - \mu_2$$

والانحراف المعياري :

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

وعلى ذلك يمكن كتابة :

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1 - \alpha,$$

حيث أن :

$$Z = \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

وحيث $z_{\frac{\alpha}{2}}$ هي قيمة لمنحنى التوزيع الطبيعي القياسي والتي تكون المساحة على يمينها تساوي $\frac{\alpha}{2}$

$\frac{\alpha}{2}$. بالتعويض عن Z بقيمتها يمكن كتابة :

$$P \left[-z_{\frac{\alpha}{2}} < \frac{(\hat{P}_1 - \hat{P}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} < z_{\frac{\alpha}{2}} \right] = 1 - \alpha .$$

وبضرب كل حد في المتباينة في $\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$ وطرح $\hat{P}_1 - \hat{P}_2$ والضرب في -1 نحصل على :

$$P \left[(\hat{P}_1 - \hat{P}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} < p_1 - p_2 \right. \\ \left. < (\hat{P}_1 - \hat{P}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \right] = 1 - \alpha .$$

إذا كانت كلا من n_2, n_1 كبيرة فإنه يمكن استبدال $\hat{p}_1 - \hat{p}_2$ بتقديرهما

$$\hat{p}_1 = \frac{x_1}{n_1}, \hat{p}_2 = \frac{x_2}{n_2} \text{ وعلى ذلك فإن :}$$

$$P \left[(\hat{P}_1 - \hat{P}_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} < (p_1 - p_2) \right. \\ \left. < (\hat{P}_1 - \hat{P}_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} \right] \approx 1 - \alpha .$$

وبالتالي يمكن الحصول على $100(1-\alpha)\%$ فترة ثقة للفرق بين نسبتي كالتالي :

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} < (p_1 - p_2)$$

$$< (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}.$$

مثال (٨ - ١١) في عينة من 2000 من الرجال و 5000 من النساء الذين يشاهدون برنامجا تلفزيونيا يوميا وجد أن 1100 من الرجال و 2300 من النساء يفضلون هذا البرنامج. أوجد 95% فترة ثقة للفرق بين نسبة كل من الرجال ونسبة كل من النساء الذين يشاهدون هذا البرنامج ويفضلونه.

الحل . بفرض أن $p_1 - p_2$ النسبتين الحقيقيتين وعلى ذلك

$$\hat{p}_2 = \frac{2300}{5000} = 0.46 \quad , \quad \hat{p}_1 = \frac{1100}{2000} = 0.55$$

وعلى ذلك التقدير بنقطة لـ $p_1 - p_2$ هو $\hat{p}_1 - \hat{p}_2 = 0.55 - 0.46 = 0.09$. باستخدام جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن $z_{0.025} = 1.96$ وبالتصويص في الصيغة التالية :

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}} < p_1 - p_2$$

$$< (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}.$$

لأن :

$$0.09 - 1.96 \sqrt{\frac{(0.55)(0.45)}{2000} + \frac{(0.46)(0.54)}{5000}} < p_1 - p_2$$

$$< 0.09 + 1.96 \sqrt{\frac{(0.55)(0.45)}{2000} + \frac{(0.46)(0.54)}{5000}},$$

والتي تختزل إلى :

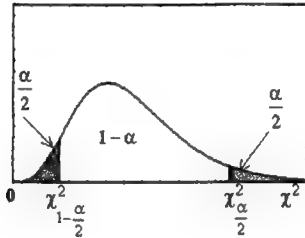
$$0.0642 < p_1 - p_2 < 0.1158.$$

Confidence Interval for the Variance (٨ - ٦) فترة ثقة للتباين

التقدير بنقطة لتباين المجتمع σ^2 نحصل عليه من تباين العينة. وعلى ذلك يعتبر S^2 مقدر للمعلمة σ^2 . التقدير بفترة للمعلمة σ^2 يمكن الحصول عليه باستخدام الإحصاء :

$$X^2 = \frac{(n-1)S^2}{\sigma^2}$$

وبالرجوع إلى نظرية (٧-١٢) فإن الإحصاء X^2 له توزيع χ^2 بدرجات حرية $n-1$ وذلك عندما نتخار العينة من توزيع طبيعي. وبالرجوع الى شكل (٨-٦).



شكل (٨-٦)

يمكن كتابة :

$$P \left[\chi^2_{1-\frac{\alpha}{2}} < X^2 < \chi^2_{\frac{\alpha}{2}} \right] = 1 - \alpha.$$

حيث $\chi^2_{\frac{\alpha}{2}}$ و $\chi^2_{1-\frac{\alpha}{2}}$ هما قيمتان لتوزيع χ^2 بدرجات حرية $n-1$ والتي المساحة على يمين $\chi^2_{\frac{\alpha}{2}}$ تساوي $\frac{\alpha}{2}$ ، والمساحة على يمين $\chi^2_{1-\frac{\alpha}{2}}$ تساوي $1 - \frac{\alpha}{2}$. بالتعويض عن X^2 بقيمتها يمكن

كتابة :

$$P \left[\chi^2_{1-\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}} \right] = 1 - \alpha.$$

بقسمة كل حد في المتباينة على $S^2 (n-1)$ ، وعكس كل حد نحصل على :

$$P \left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} \right] = 1 - \alpha.$$

لعينة عشوائية خاصة من الحجم n ، فإن تباين العينة s^2 يحسب ويمكن الحصول على $100\%(1 - \alpha)$ فترة ثقة للمعلمة σ^2 كما يلي :

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}}.$$

مثال (٨-١٢) تسلم أحد التجار كمية كبيرة من بطاريات السيارات المنتجة بواسطة مصنع جديد وتم اختيار عينة عشوائية من البطاريات التي تسلمها التاجر وتمت تجربتها فكانت أعمارها بالشهر هي :

26.9	28.5	33.6	28.0	23.9
28.7	29.3	29.1	35.9	35.2

أوجد 99% فترة ثقة للمعلمة σ^2 .

الحل . أولاً نحصل على تباين العينة s^2 وهو :

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right]$$

$$= \frac{1}{9} \left[9076.87 - \frac{(299.1)^2}{10} \right] = 14.53.$$

باستخدام جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $v = n - 1 = 9$ فإن :

$$\chi^2_{0.995} = 1.735 \quad , \quad \chi^2_{0.005} = 23.587$$

بالتعويض في الصيغة التالية :

$$\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2}}^2}$$

يمكن الحصول على فترة ثقة كالتالي :

$$\frac{(9)(14.53)}{23.587} < \sigma^2 < \frac{(9)(14.53)}{1.735}$$

والتي يمكن اختزالها الى

$$5.544 < \sigma^2 < 75.372.$$

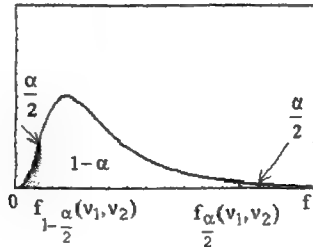
(٧-٨) فترة ثقة لنسبة تباينين

Confidence Interval for the Ratio of two variances

التقدير بنقطة لنسبة تباينين مجتمعين ، σ_1^2 / σ_2^2 ، يمكن الحصول عليه من النسبة ، S_1^2 / S_2^2 ، لتباينين عتين. وعلى ذلك يعتبر S_1^2 / S_2^2 مقدر للنسبة σ_1^2 / σ_2^2 . فإذا كان لدينا مجتمعان الأول يتبع توزيعاً طبيعياً متوسطه μ_1 وتباينه σ_1^2 والثاني يتبع توزيعاً طبيعياً متوسطه μ_2 وتباينه σ_2^2 ، فإنه يمكن الحصول على $100\%(1-\alpha)$ فترة ثقة لـ σ_1^2 / σ_2^2 باستخدام الإحصاء :

$$F = \frac{\sigma_2^2}{\sigma_1^2} \frac{S_1^2}{S_2^2}.$$

تبعاً لنظرية (٧-١٣) فإن المتغير العشوائي F له توزيع F بدرجات حرية $v_1 = n_1 - 1$ ، $v_2 = n_2 - 1$. وعلى ذلك يمكن كتابة (كما يتضح من شكل (٧-٨) .



شكل (٧-٨)

$$P \left[f_{1-\frac{\alpha}{2}}(v_1, v_2) < F < f_{\frac{\alpha}{2}}(v_1, v_2) \right] = 1 - \alpha.$$

حيث $f_{\frac{\alpha}{2}}(v_1, v_2)$, $f_{1-\frac{\alpha}{2}}(v_1, v_2)$ هما قيمتان لتوزيع F بدرجات حرية v_1 , v_2

على التوالي (كما يتضح من شكل (٧-٨)). وبالتعويض عن F بقيمتها فإن :

$$P \left[f_{1-\frac{\alpha}{2}}(v_1, v_2) < \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} < f_{\frac{\alpha}{2}}(v_1, v_2) \right] = 1 - \alpha.$$

بضرب كل حد في المتباينة في S_2^2 / S_1^2 وعكس كل حد نحصل على :

$$P \left[\frac{S_1^2}{S_2^2} \frac{1}{f_{\frac{\alpha}{2}}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} \frac{1}{f_{1-\frac{\alpha}{2}}(v_1, v_2)} \right] = 1 - \alpha.$$

نتائج نظرية (٧-١٤) سوف تساعدنا في استبدال $f_{1-\frac{\alpha}{2}}(v_1, v_2) \rightarrow f_{\frac{\alpha}{2}}(v_2, v_1)$

وعلى ذلك :

$$P \left[\frac{S_1^2}{S_2^2} \frac{1}{f_{\frac{\alpha}{2}}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} \frac{1}{f_{\frac{\alpha}{2}}(v_2, v_1)} \right] = 1 - \alpha.$$

لأي عتين عشوائيتين مستقلتين من الحجم n_2, n_1 مأخوذتين من مجتمعين طبيعيين ،

فإن النسبة S_1^2 / S_2^2 تحسب ويتم الحصول على 100% (1- α) فترة ثقة كما يلي :

$$\frac{S_1^2}{S_2^2} \frac{1}{f_{\frac{\alpha}{2}}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} \frac{1}{f_{\frac{\alpha}{2}}(v_2, v_1)}.$$

مثال (٨-١٣) إذا كانت درجات كل من الطلاب والطالبات بسأحدى الجامعات في مادة الإحصاء يتبع توزيعاً طبيعياً. اختبر عينة عشوائية من بين الطلاب وأخرى من بين الطالبات فكانت درجاتهم كما يلي :

الطلاب :	59	79	73	49	88	83	69	44	81
الطالبات :		74	79	49	59	82	69	79	89

أوجد 90% فترة ثقة للنسبة σ_1^2 / σ_2^2 .

الحل .

$$n_1 = 9, s_1 = 15.57, n_2 = 8, s_2 = 13.07$$

$f_{0.05}(7,8)=3.5, f_{0.05}(8,7)=3.73$ المستخرجتان من جدول توزيع F في ملحق (٦)

بدرجات حرية $v_1 = 8, v_2 = 7$ للعينات الأولى ودرجات حرية $v_1 = 7, v_2 = 8$ للعينات

الثانية. يمكن الحصول على 90% فترة ثقة للنسبة σ_1^2 / σ_2^2 وذلك بالتعويض في الصيغة التالية

$$\frac{s_1^2}{s_2^2} \frac{1}{f_{\frac{\alpha}{2}}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\frac{\alpha}{2}}(v_2, v_1) .$$

أي أن :

$$\frac{(15.57)^2}{(13.07)^2 (3.73)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{(15.57)^2 (3.5)}{(13.07)^2}$$

والتي تختزل إلى :

$$0.3805 < \frac{\sigma_1^2}{\sigma_2^2} < 4.967.$$

تمارين :

١- إذا كان عمر المصابيح الكهربائية المنتجة في مصنع ما ، تقريباً تتبع التوزيع الطبيعي بانحراف معياري 40 ساعة . اختبرت عينة عشوائية من 25 مصباحاً ووجد أن متوسط العمر 780 ساعة. أوجد 99% فترة ثقة لمتوسط المجتمع لكل المصابيح المنتجة من هذا المصنع .

٢- في مركز تجاري كبير ، صممت ماكينة للعصير بحيث أن كمية العصير المستخرجة منها لكل كوب يتبع تقريباً توزيعاً طبيعياً بانحراف معياري 0.5 أوقية لكل كوب . أوجد 95% فترة ثقة للمتوسط μ إذا اختبرت عينة عشوائية من 26 كوب ووجد أن $\bar{x} = 7.4$ أوقية.

٣- إذا كانت الأطوال لعينة عشوائية من 50 طالباً لها متوسط $\bar{x} = 68.5$ بوصة وانحراف معياري $s = 2.7$ بوصة. المطلوب

(أ) إيجاد 99% فترة ثقة للمتوسط μ .

- (ب) تقدير حجم العينة n حتى يمكننا التأكد باحتمال 0.99 من أن الخطأ لا يتجاوز 0.5 .
- ٤- في دراسة عن تلوث الهواء بأكسيد الكبريت المنبعث من إحدى المصانع. اختبرت عينة عشوائية من قراءات 70 يومياً وحسب متوسط العينة فوجد أن يساوي 19.1 طنساً بانحراف معياري قدره 5.22 أطنان . أوجد 95% فترة ثقة للمتوسط μ .
- ٥- يرغب صاحب مصنع في تقدير حجم العينة n حتى يمكنه التأكد تأكداً معقولاً من أن تقديره وباحتمال 0.99 لن يكون مخطئاً بأكثر من 4 وحدات معينة إذا علم أن الانحراف المعياري يساوي 18 وحدة. أوجد حجم العينة التي تحقق الشروط التي وضعاها صاحب المصنع.
- ٦- مونت 35 سيارة بكمية قدرها 4.8 جالون من البنزين المضاف إليه مادة معينة وسارت حتى نفذ الوقود. وبعد تمام التجربة حسب متوسط الأميال التي قطعتها كل سيارة بكل جالون فكان $\bar{x} = 18.9$ لكل جالون بانحراف معياري $s = 2.4$. أوجد 95% فترة ثقة للمتوسط μ .
- ٧- مجموعة قوامها 50 من حيوانات التجارب أطعمت نوعاً معيناً من المقتنات الغذائية لمدة شهر فكان متوسط الزيادة في الوزن $\bar{x} = 30$ بانحراف معياري $s = 5$ أولية أوجد 99% فترة ثقة للمتوسط μ .
- ٨- دلت الخبرة مع العمال المشغلين في صناعة معينة أن الزمن الذي يحتاج إليه العامل لإكمال عمل يتبع توزيعاً طبيعياً بانحراف معياري 20 دقيقة فإذا اختبرت عينة عشوائية من 25 عاملاً ووجد أن $\bar{x} = 13$ دقائق أوجد 95% فترة ثقة للمتوسط μ .
- ٩- إذا كان ضغط الدم لمجموعة من الأفراد يتبع التوزيع الطبيعي بانحراف معياري $\sigma = 2$ اختبرت عينة عشوائية حجمها 25 فرداً من هذه المجموعة وكان متوسط ضغط الدم من العينة $\bar{x} = 91$ أوجد 99% فترة ثقة للمتوسط μ .
- ١٠- اختبرت عينة عشوائية من 100 أسرة فكان متوسط دخل الأسرة في العينة $\bar{x} = 1000$ جنيهاً بانحراف معياري $s = 100$ جنيهاً. أوجد 95% فترة ثقة للمتوسط μ .
- ١١- اختبرت عينة عشوائية من 100 مريض بالسكر وكان متوسط أعمارهم $\bar{x} = 55$ بانحراف معياري $s = 20$ أوجد 99% فترة ثقة للمتوسط μ .
- ١٢- للدراسة النمو لنوع معين من الزهور اختبرت عينة عشوائية من 70 زهرة ووجد أن متوسط النمو خلال عام $\bar{x} = 40.9$. بانحراف معياري $s = 5.1$. أوجد 95% فترة ثقة للمتوسط μ .

١٣- في بحث ميداني ، اختبرت عينة عشوائية من 400 عائلة ووجد أن متوسط ما أنفق على الطعام خلال عام هو 1200 جنية بانحراف معياري قدره 100 جنيهاً أوجد 99% لفترة ثقة للمتوسط μ .

١٤- بفرض أن أوزان الدببة في لعب الأطفال يتبع التوزيع الطبيعي بانحراف معياري 3.5 كيلو جرام ما هو حجم العينة اللازم بدقة 95% أن متوسط العينة لا يختلف عن متوسط المجتمع بأكثر من 0.5 كيلو جرام .

١٥- إذا كان أطوال الطلبة المستجدين في كلية ما يتبع توزيعاً طبيعياً بانحراف معياري 2.5 بوصة. فإذا اختبرت عينة عشوائية من الحجم $n = 25$ من هؤلاء الطلبة ووجد أن $\bar{x} = 68.5$ بوصة أوجد 95% فترة ثقة للمتوسط μ .

١٦- في دراسة عن حجم المبيعات اليومية في محل تجاري تم تسجيل حجم المبيعات خلال 60 يوماً ، خلال فترة قدرها سنة ، وتم حساب متوسط المبيعات اليومية فكانت $\bar{x} = 218$ جنيهاً بانحراف معياري $s = 18$ جنيهاً. المطلوب تقدير 99% فترة ثقة للمتوسط μ .

١٧- إذا كانت أسعار إحدى السلع تتبع توزيعاً طبيعياً بانحراف معياري 4 جنيهاً. اختبرت عينة عشوائية من أسعار هذه السلعة حجمها $n = 15$ فكان وسطها الحسابي $\bar{x} = 40$ جنيهاً. أوجد 99% ، 95% فترة ثقة للمتوسط μ .

١٨- إذا كان معروفاً أن الضغط الداخلي لكرات التنس المنتجة بواسطة أحد المصانع يتبع توزيعاً طبيعياً بانحراف معياري $\sigma = 0.25$ رطلاً لكل بوصة. اختبرت عينة عشوائية من الحجم $n = 28$ من الكرات المنتجة من هذا المصنع وكان $\bar{x} = 28$ رطلاً لكل بوصة. أوجد 99% فترة ثقة للمتوسط μ .

١٩- اختبرت عينة عشوائية من الحجم $n=50$ من إنتاج آلة ملاصقة الأرز في أكياس فحصلنا على البيانات التالية $\bar{x} = 4.80$ كجم و $s = 0.6$ كجم. أوجد 95% فترة ثقة للمتوسط μ .

٢٠- اختبرت عينة عشوائية من 100 فاتورة مبيعة وذلك من مجتمع كبير جداً من الفواتير المبيعة. فإذا كان متوسط العينة هو $\bar{x} = 18.5$ جنيهاً بانحراف معياري 6.0 جنيهاً. أوجد 99% فترة ثقة للمتوسط μ .

٢١- اختبرت عينة عشوائية من الفواتير المبيعة من الحجم $n = 100$ وتم وضعها في التوزيع التكراري التالي :

حدود الفئة	1-50	51-100	101-150	151-200	201-250	251-300
------------	------	--------	---------	---------	---------	---------

عدد القواتير	10	19	15	26	13	17
-----------------	----	----	----	----	----	----

استخدم المعلومات في التوزيع التكراري السابقة في إيجاد 95% و 99% فترة ثقة للمتوسط المجتمع.

-٢٢- إذا كانت أوزان 10 صناديق من القمح هي :

12, 10.2, 10.1, 9, 10.1, 10.2, 10.1, 9.8, 9.9, 8.9

(الأوزان مقاسة بالأوقية). أوجد 95% فترة ثقة للمتوسط μ وذلك تحت فرض أن وزن الصندوق من القمح يتبع توزيعاً طبيعياً .

-٢٣- للدراسة اختفاء العلامات الأرضية البيضاء وسط الطرق نتيجة المرور الكثيف. أخذت عينة من 8 مناطق مختلفة ولوحظ اختفاء العلامات بعد مرور السيارات مقربة لأقرب مائة سيارة كالتالي :

163800, 136400, 108200, 125400,
143700, 163000, 159400, 122600

أوجد 99% فترة ثقة للمتوسط μ وذلك تحت فرض أن توزيع المجتمع الذي اختبرت منه العينة يتبع توزيعاً طبيعياً .

-٢٤- أخذت عينة عشوائية من 20 آلة حاسبة من نوع ما ووجد أن متوسط أعمارهم $\bar{x} = 8$ سنوات بانحراف معياري $s = 2.6$ سنة . أوجد 95% فترة ثقة للمتوسط μ تحت فرض أن توزيع المجتمع الذي اختبرت منه العينة طبيعياً .

-٢٥- إذا كانت كمية النيكوتين في 7 سجاير من نوع معين مقاس بالمليجرامات هو

21, 19, 23, 19, 23, 18, 17

أوجد 95% فترة ثقة للمتوسط μ تحت فرض أن كمية النيكوتين في هذا النوع من السجاير تتبع توزيعاً طبيعياً.

-٢٦- في عينة عشوائية من 10 مرضى في العناية المركزة في مستشفى ما وجد أن متوسط درجة حرارة الجسم 38.1 درجة بانحراف معياري 1.2 . أوجد 95% فترة ثقة للمتوسط المجتمع μ تحت فرض أن متوسط درجة حرارة الجسم للمرضى في العناية المركزة يتبع توزيعاً طبيعياً.

-٢٧- في إحدى المراكز الصحية أجريت دراسة على عينة ميدلية ، على الكمية المتوقعة من الأكسجين (بالتر في الدقيقة) الذي يستهلكه الفرد الذي عمره 20-25 سنة ووجد أن الانحراف المعياري 0.5 لتر في الدقيقة. المطلوب تقدير حجم العينة اللازم لبحث قادم حتى يمكننا التأكد باحتمال 0.99 من أن الخطأ لا يتجاوز 0.1 لتر في الدقيقة .

-٢٨- إذا كانت أوزان طلاب إحدى المدارس تتبع توزيعاً طبيعياً. اختبرت عينة عشوائية حجمها 16 طالباً من هذه المدرسة فكانت أوزانهم بالكيلو جرام كما يلي :

42.1	44.2	49	44	45	44	51	52
45.2	46	51.3	44	47	49	41.2	47

أوجد 95% فترة ثقة للمتوسط μ .

-٢٩- في تجربة على 10 من رواد الفضاء في مجال يحاكي مجال انعدام الوزن وجد أن متوسط ضربات القلب لم 26.22 دقة في الدقيقة بانحراف معياري 4.3 دقة في الدقيقة. أوجد 99% فترة ثقة للمتوسط μ تحت فرض أن ضربات القلب لرواد الفضاء تتبع توزيعاً طبيعياً .

-٣٠- في دراسة للسعرات الحرارية المنتجة من نوع معين من القمح تم الحصول على البيانات التالية من عينة عشوائية حجمها $n=5$:

8000, 7820, 8200, 8470, 8123

أوجد 95% فترة ثقة للمتوسط μ تحت فرض أن السعرات تحت الدراسة تتبع توزيعاً طبيعياً .

-٣١- اختبرت عينة عشوائية من 10 كرات وتم قياس قطر كل كرة وحساب متوسط القطر فكان $\bar{x} = 4.9$ ملم بانحراف معياري 0.07 ملم. أوجد 95% فترة ثقة للمتوسط μ (تحت فرض أن العينة تم اختيارها من مجتمع طبيعي).

-٣٢- لمقارنة صنفين من القمح من حيث كمية الغصول أخذت 5 فدادين لكل صنف من القمح وزرع فيها الصنفين تحت نفس الظروف. أعطى الصنف A في المتوسط 78.3 وحدة ما لكل فدان بانحراف معياري 5.5 وحدة لكل فدان. بينما أعطى الصنف B في المتوسط 88 وحدة لكل فدان بانحراف معياري 6.2 أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$ تحت فرض أن العنيتين تم اختيارهما من مجتمعين طبيعيين.

-٣٣- في دراسة لتقدير الفرق بين الأيجور في كليتين A و B أخذت عينة عشوائية من $n_1=25$ أستاذ من الكلية A ووجد أن متوسط الأجر خلال 9 شهور هو 10000 دولار بانحراف معياري 1100 دولار. كما أخذت عينة عشوائية أخرى من الحجم $n_2=20$ أستاذ من الكلية B ووجد أن متوسط الأجر خلال 9 شهور 13000 دولار بانحراف معياري 1120 دولار. أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$ تحت فرض أن $\sigma_1^2 = \sigma_2^2$ وأن العنيتين تم اختيارهما من توزيعين طبيعيين.

-٣٤- تحاول شركة لتأجير السيارات اتخاذ قرار بشأن شراء إطارات من النوع A أو من النوع B. لتقدير الفرق بين النوعين أجرت تجربة حيث استخدمت 12 إطاراً من كل نوع وجربت الإطارات حتى انتهاء عمرها وكانت النتائج كما يلي (المسافة التي قطعها السيارة بالأميال):

A النوع الأول : $\bar{x}_1 = 22500$, $s_1 = 310$

B النوع الثاني : $\bar{x}_2 = 23600$, $s_2 = 380$

بفرض أن العنيتين تتم اختيارهما من توزيعين طبيعيين وأن $\sigma_1^2 \neq \sigma_2^2$ أوجد 95% فترة ثقة لـ $\mu_1 - \mu_2$.

٣٥- البيانات التالية تمثل الزمن اللازم لعرض فيلم منتج من قبل شركتين مختلفتين .

	الزمن (بالدقائق)					
الشركة الأولى	100	90	110	80	90	
الشركة الثانية	96	120	90	174	80	108 107

أحسب 90% فترة ثقة لـ $\mu_1 - \mu_2$ بفرض أن أزمدة العرض تقريباً تتبع التوزيع الطبيعي.

٣٦- أجريت دراسة على طريقة جديدة لتخفيض الوزن 10 أرطال في المتوسط على 7 سيدات تابعا نفس الطريقة لتخفيض الوزن وسجلت الأوزان قبل وبعد أسبوعين

السيدة	1	2	3	4	5	6	7
الوزن قبل	128	132	136	150	140	137	124
الوزن بعد	129	120	127	135	128	130	159

أوجد 99% فترة ثقة لـ μ_D إذا علم أن توزيعات الأوزان تقريباً طبيعية .

٣٧- مقارنة عنيقتين من ناحية تأثيرهما على نمو العجول خلال شهرين من التغذية اختصرت عينة عشوائية من 10 أزواج (توائم من المجول) لكل زوج أعطى إحدى التوأمين العليقة A و الآخر العليقة B وتم الحصول على النتائج التالية : $\bar{d} = 10$, $s_{\bar{d}} = 3$ أوجد 95% فترة ثقة للمعلمة μ_D .

٣٨- طبق مقياس للاتجاهات نحو الأطفال في عينة من السيدات الآتي تم قبولهم بقسم دراسات الطفولة فور التحاقهن بالمقسم ، ثم طبق نفس المقياس مرة أخرى فور حصولهن على البكالوريوس وتم الحصول على النتائج التالية $\bar{d} = 18$, $s_{\bar{d}} = 3$, $n = 10$ أوجد 99% فترة ثقة للمعلمة μ_D .

٣٩- قام شخص بإجراء 10 عمليات حسابية على آلتين حاسبتين وتسجيل الزمن اللازم لكل عملية على كل آلة وقد تم الحصول على البيانات التالية $\bar{d} = 9$, $s_{\bar{d}} = 4$ أوجد 99% فترة ثقة للمعلمة μ_D .

٤٠- في عينة عشوائية من 1000 شخص من القراء لصحيفة ما وجد أن 400 منهم يفضلون قراءة نوع معين من الإعلانات في الصحيفة أوجد 99% فترة ثقة للنسبة p .

٤١- في مباراة رياضية للجري وجد أن 240 طالب من 400 طالب يمكنهم الجري لمدة ميل في أقل من 7 دقائق . أوجد 95% فترة ثقة للنسبة p .

٤٢- أراد أحد مكاتب الاستقصاء معرفة نسبة الأصوات التي تؤيد مرشحاً معيناً في الانتخابات. اختبرت عينة عشوائية من 100 شخص فوجد أن من بينهم 55 فرداً يؤيدون هذا المرشح أوجد 99% فترة ثقة للنسبة p .

٤٣- أخذت عينة عشوائية من 200 طالب فكان عدد الناجحين منهم 144 طالب . المطلوب (أ) نسبة النجاح في العينة

(ب) 95% فترة ثقة لنسبة النجاح p .

٤٤- قدر حجم العينة العشوائية التي يمكن اختيارها لتقدير نسبة الطلاب الناجحين في مادة الإحصاء إذا كانت نسبة النجاح في عينة مبدئية 0.75 وبشرط أن أقصى خطأ لا يزيد عن 0.05 بنقبة 95% .

٤٥- في إحدى مؤسسات رعاية الأحداث بما 6000 نزيل ، أخذت عينة عشوائية مكونة من 300 وأجريت دراسة لتحديد سبب دخول الحدث المؤسسة فوجد أن 80% من العينة يرجع سبب الدخول للمؤسسة إلى عدم رعاية الأم له. أوجد 99% فترة ثقة لنسبة الأحداث في المؤسسة الذين يرجع سبب دخولهم المؤسسة إلى عدم رعاية الأم.

٤٦- أخذت عينة عشوائية من 100 شخص من أعضاء هيئة التدريس ووجد أن 20% منهم أجورهم تزيد عن 30000 دولار في السنة. أوجد فترة ثقة للنسبة p .

٤٧- اتحاد جامعي يرغب في تقدير نسبة الطلبة الذين يفضلون هيئة طلابية جديدة. اختبرت عينة عشوائية من 400 طالب. فإذا أعطت نتائج التصويت $\hat{p} = 0.8$. أوجد 95% فترة ثقة للنسبة p .

٤٨- اختبرت عينة عشوائية من 400 مواطن في مجتمع سكاني ما ووجد أن 300 منهم يفضلون إضافة قليل من الفلور إلى مياههم. استخدم هذه البيانات في إيجاد 99% فترة ثقة للنسبة p .

٤٩- اختبرت عينة عشوائية من 600 مدخن سجاير في مجتمع سكاني ما ووجد أن 80 منهم يفضلون النوع A استخدم هذه البيانات في إيجاد :

(أ) نسبة النجاح في العينة

(ب) 95% فترة ثقة لنسبة النجاح p .

(ت) تقدير حجم العينة التي يمكن اختيارها لتقدير نسبة المدخنين الذين يفضلون النوع A

وذلك بثقة 95% إذا كان حجم الخطأ 0.172.

٥٠- في دراسة لنسبة ربات البيوت اللاتي يمتلكن غسالة بمجفف وجد أن 55 من 100

سيدة في المدينة A يمتلكن غسالة بمجفف بينما في المدينة B وجد أن 45 من 150 يمتلكن

غسالة بمجفف. أوجد 95% فترة ثقة لـ $p_1 - p_2$.

٥١- أوجد 95% فترة ثقة للمعلمة σ^2 للمثال 19.

٥٢- أوجد 95% فترة ثقة للمعلمة σ^2 للمثال 20.

٥٣- أوجد 99% فترة ثقة للمعلمة $\frac{\sigma_1^2}{\sigma_2^2}$ للمثال 32.

٥٤- أوجد 95% فترة ثقة للمعلمة $\frac{\sigma_1^2}{\sigma_2^2}$ للمثال 33.

الفصل التاسع

اختبارات الفروض

Tests of Hypotheses

(٩-١) الفروض الإحصائية Statistical Hypotheses

تعتبر اختبارات الفروض الإحصائية أهم فرع في نظرية القرارات. أولاً، دعنا نعرف بدقة ماذا نعني بالفرض الإحصائي .

تعريف : الفرض الإحصائي هو جملة ما تخص واحد أو أكثر من المجتمعات، من الممكن أن تكون صحيحة أو غير صحيحة .

للتأكد من صحة أو عدم صحة الفرض الإحصائي لا بد من دراسة كل مفردات المجتمع تحت الدراسة وهذا بالطبع غير عملي في معظم الحالات. بدلاً من ذلك فإننا نختار عينة عشوائية من المجتمع ونستخدم المعلومات الموجودة في العينة لتتخذ قرار بقبول أو رفض الفرض الإحصائي. القرار الذي نتخذه سوف يكون سليم إذا كان الفرض صحيح وتم قبوله أو خطأ وتم رفضه. بينما يكون القرار غير سليم إذا كان الفرض صحيح وتم رفضه أو غير صحيح وتم قبوله.

الفروض التي نضعها على أمل أن نرفضها تسمى فروض العدم H_0 . ويرمز لفرض العدم بالرمز H_0 . رفض فرض العدم يؤدي إلى قبول فرض بديل H_1 hypothesis alternative ويرمز للفرض البديل بالرمز H_1 . فعلى سبيل المثال إذا كان فرض العدم H_0 أن متوسط الطول في مجتمع ما $\mu = 160$ (مقاسه بالسنتيمتر) فإن الفرض البديل H_1 قد يكون $\mu \neq 160$ أو $\mu < 160$ أو $\mu > 160$.

(٩-٢) الخطأ من النوع الأول والخطأ من النوع الثاني :

Type I Error and Type II Error

سوف نسهل المفاهيم المستخدمة في اختبارات الفروض والتي تخص مجتمع ما بالمثال التالي : بفرض أنه تم إجراء اختبار قدرات لعدد من المتقدمين لشغل وظيفة ما في مجال الحاسب الآلي. يشتمل الاختبار على 15 سؤال وكل سؤال له 5 أجوبة ممكنة واحد منهم الصحيح. من شروط اجتياز الاختبار حصول المتقدم على 7 درجات فأكثر. هذا يعني أن المتقدم لديه بعض المعلومات والتي تؤهله للعمل في مجال الحاسب الآلي. فرض العدم في هذه الحالة أن معلومة ذي الحدين (احتمال النجاح) في محاولة معطاة (سؤال) هو $p = \frac{1}{5}$ ، أى أن الشخص المقدم للاختبار يعتمد على التخمين. الفرض البديل في هذه الحالة $p > \frac{1}{5}$. وعلى ذلك يمكن وضع فرض العدم والفرض البديل على الصورة التالية :

$$H_0 : p = \frac{1}{5}$$

$$H_1 : p > \frac{1}{5}$$

أن الطريقة السابقة في اتخاذ القرار قد تؤدي إلى استنتاجين غير صحيحين . فقد يحصل المتقدم للوظيفة على 7 درجات أو أكثر عن طريق التخمين. في هذه الحالة نكون قد وقعنا في خطأ عند رفض فرض العدم وقبول الفرض البديل مع أن H_0 صحيحا. مثل هذا الخطأ يسمى خطأ من النوع الأول.

تعريف : يحدث الخطأ من النوع الأول إذا كان فرض العدم صحيح ويتخذ قرار برفضه .
النوع الثاني من الخطأ والذي يمكن الوقوع فيه إذا حصل المتقدم للوظيفة على درجة أقل من 7 درجات وتستنتج أنه يضمن بينما هو في الحقيقة لديه بعض المعلومات عن الإجابة الصحيحة.

تعريف : يحدث الخطأ من النوع الثاني إذا قبلنا فرض العدم وهو خطأ.
يعتمد قرارنا في هذا المثال على الإحصاء X الذي يمثل عدد الإجابات الصحيحة التي يحصل عليها المتقدم في الاختبار حيث أن $x = 0, 1, 2, \dots, 15$. القيم الممكنة من 0 إلى 15 تقسم إلى مجموعتين. تضم المجموعة الأولى القيم الأقل من 7 أما المجموعة الثانية فتضم القيم التي تساوي 7 فأكثر. كل الدرجات الممكنة والتي تساوي 7 فأكثر تكون منطقة الرفض $region rejection$ بينما الدرجات التي أقل من 7 تكون منطقة القبول $acceptance region$. الرقم 7 يسمى القيمة الحرجة $critical region$. إذا وقعت قيمة الإحصاء الذي يعتمد عليه قرارنا في منطقة الرفض نرفض فرض العدم H_0 ونقبل الفرض البديل H_1 بينما إذا وقعت قيمة الإحصاء في منطقة القبول نقبل H_0 ونرفض H_1 .

تعريف : احتمال الوقوع في خطأ من النوع الأول يسمى مستوى المنوية $level of significance$ للاختبار ويرمز له بالرمز α .

في مثالنا فإن الخطأ من النوع الأول يقع إذا حصل المتقدم للاختبار على 7 درجات أو أكثر عن طريق التخمين. فإذا كانت X تمثل عدد الإجابات الصحيحة فإن :

$$\alpha = P(\text{الخطأ من النوع الأول})$$

$$= P(\text{رفض } H_0 \text{ وهو صحيح})$$

$$= P(X \geq 7 | p = \frac{1}{5})$$

$$= \sum_{x=7}^{15} b \left(x; 15, \frac{1}{5} \right)$$

$$= 1 - \sum_{x=0}^6 b \left(x; 15, \frac{1}{5} \right)$$

$$\approx 1 - 0.982 = 0.018.$$

وهذا يعنى أنه تقريباً في 1.8% من كل التجارب من هذا النوع حيث $n = 15$ تؤدي إلى الوقوع في خطأ من النوع الأول، أى رفض H_0 وهو صحيحاً. في هذه الحالة يمكن القول أن $H_0 : p = \frac{1}{5}$ نختار عند مستوى معنوية $\alpha = 0.018$. أحياناً تسمى α حجم منطقة الرفض . size of rejection

احتمال الوقوع في خطأ من النوع الثاني يرمز له بالرمز β ولحساب قيمة β لا بد من وضع فرض بديل معين. فمثلاً عند اختبار فرض العدم $p = \frac{1}{5}$ ضد الفرض البديل

$p = \frac{7}{10}$ فإننا نكون قادرين على حساب احتمال حصول المتقدم للاختبار على درجة أقل من 7. عندما تكون $p = \frac{7}{10}$ في هذه الحالة :

$$\beta = P(\text{خطأ من النوع الثاني})$$

$$= P(\text{قبول } H_0 \text{ وهو خطأ})$$

$$= P(X < 7 \mid p = \frac{7}{10})$$

$$= \sum_{x=0}^6 b \left(x; 15, \frac{7}{10} \right) = 0.015$$

وهذا يعنى أن تقريباً 1.5% من كل التجارب من هذا النوع حيث $n = 15$ تؤدي إلى قبول فرض العدم وهو خطأ. في الحقيقة لكل قيمة من α لا يوجد قيمة وحيدة من β بل يوجد قيم مختلفة لـ β وذلك لقيمة $p = \frac{1}{5}$. على سبيل المثال يوجد قيمة لـ β عند $p = 0.6$ وقيمة لـ β عند $p = 0.7$ وهكذا. الجدول التالي يعطى قيم β لقيم مختلفة مختارة من p (كل واحدة تمثل الفرض البديل لفرض العدم $H_0 : p = \frac{1}{5}$).

p	0.6	0.7	0.8	0.9
β	0.095	0.015	0.001	0.000

يلاحظ من الجدول أنه كلما بعدت قيمة p (تحت الفرض البديل) عن $p = \frac{1}{5}$ كلما قل

الوقوع في خطأ من النوع الثاني أي كلما قلت قيمة β عند منطقة الرفض $X \geq 7$.

مثال (٩-١) إذا كانت p تمثل نسبة الأصوات المؤيدة للشخص A ضد الشخص B في ترشيح ما. اختيرت عينة عشوائية من الحجم $n = 20$ فإذا كانت X تمثل عدد الأشخاص الذين يؤيدون الشخص A في العينة.

المطلوب : (أ) حساب احتمال الوقوع في خطأ من النوع الأول إذا كان $H_0: p = 0.5$ ضد الفرض البديل $H_1: p = 0.5$ ومنطقة الرفض $X \leq 5$ أو $X \geq 12$.

(ب) حساب احتمال الوقوع في الخطأ من النوع الثاني تحت الفرض البديل

$$H_1: p = 0.9$$

$$\alpha = P(\text{الخطأ من النوع الأول}) \quad (أ)$$

$$= P(\text{رفض } H_0 \text{ وهو صحيح})$$

$$= P(X \geq 12 | p = 0.5)$$

$$+ P(X \leq 5 | p = 0.5)$$

$$= (1 - \sum_{x=0}^{11} b(x; 20, 0.5))$$

$$+ \sum_{x=0}^5 b(x; 20, 0.5)$$

$$= (1 - 0.748) + 0.021$$

$$= 0.252 + 0.021 = 0.273$$

$$\beta = P(\text{الخطأ من النوع الثاني}) \quad (ب)$$

$$= P(\text{قبول } H_0 \text{ وهو خطأ})$$

$$= P(5 < X < 12 | p = 0.9) = P(6 \leq X \leq 11 | p = 0.9)$$

$$= \sum_{x=0}^{11} b(x; 20, 0.9) - \sum_{x=0}^5 b(x; 20, 0.9)$$

$$= 0.0000 - 0.0000 = 0.0000.$$

عموماً الاختبار الجيد هو الذي يجعل كلا من α , β أصغر ما يمكن ومن الصعب الحصول على هذا الاختبار لأنه لا يمكن تصغير كل من α , β في آن واحد حيث أن تصغير أحدهما يؤدي إلى تكبير الأخرى. لذلك لجأ الإحصائيون إلى تثبيت مستوى المعنوية α عند قيمة محددة ثم اختيار الاختبار الإحصائي الذي يجعل β أقل ما يمكن. من القيم الشائعة لمستوى المعنوية $\alpha = 0.05$ أو $\alpha = 0.01$. يقال للاختبار أنه معنوي significant عند $\alpha = 0.05$ إذا رفض فرض العدم

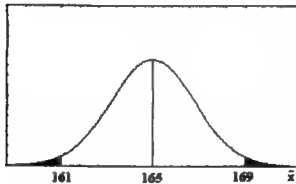
عند مستوى معنوية $\alpha = 0.05$ و يعتبر الاختبار معنوي جداً إذا رفض فرض العدم عند مسوي معنوية $\alpha = 0.01$.

من السهل توضيح المفاهيم السابقة بياناً عندما يكون المجتمع متصلاً . على سبيل المثال إذا كان فرض العدم أن متوسط الطول لمجموعة من الطلبة في جامعة ما هو $\mu = 165$ سم ضد الفرض البديل أن متوسط الطول لا يساوي 165 سم . أى أننا نرغب في اختبار :

$$H_0 : \mu = 165,$$

$$H_1 : \mu \neq 165.$$

الفرض البديل يمكن أن يكون $\mu < 165$ أو $\mu > 165$. بفرض أن الانحراف المعياري لمجموع الأطوال $\sigma = 10$ ، الإحصاء الذي سوف نبني عليه قرارنا والذي يعتمد على عينة عشوائية من الحجم $n = 30$ سوف يكون \bar{X} والذي يعتبر التقدير الأكثر كفاءة للمعلمة μ . علمنا في الفصل السابع أن التوزيع العيني للإحصاء \bar{X} تقريباً يتبع التوزيع الطبيعي بانحراف معياري $\sigma_{\bar{X}} = \sigma / \sqrt{n} = 10 / \sqrt{30} = 1.8$. بفرض أننا اخترنا منطقة الرفض ممثلة في قيم \bar{X} الأقل من 161 أو قيم \bar{X} التي أكبر من 169 . أى أن $\bar{X} < 161$ أو $\bar{X} > 169$ والموضحة في شكل (٩-١) . منطقة القبول سوف تكون $161 < \bar{X} < 169$. وعلى ذلك إذا وقع متوسط العينة \bar{X} في منطقة القبول نقبل H_0 ونرفض H_1 وغير ذلك نرفض H_0 . احتمال الوقوع في خطأ من النوع الأول، أو مستوى المعنوية للاختبار، يساوي مجموع المساحات المظلمة في كل من جانبي التوزيع العيني للإحصاء \bar{X} في شكل (٩-١) .



شكل (٩-١)

وعلى ذلك فإن :

$$\alpha = P(\bar{X} < 161 \mid \text{صحيح } H_0) \\ + P(\bar{X} > 169 \mid \text{صحيح } H_0)$$

القيمتين الحرجتين للمتغير العشوائي Z والمقابلتين للقيمتين $\bar{x}_2 = 169$, $\bar{x}_1 = 161$ عندما تكون H_0 صحيحة هما :

$$z_1 = \frac{161 - 165}{1.8} = -2.22, \\ z_2 = \frac{169 - 165}{1.8} = 2.22.$$

وعلى ذلك فإن :

$$\alpha = P(Z < -2.22) + P(Z > 2.22) \\ = 2 P(Z > 2.22) \\ = 2 [0.5 - P(0 < Z < 2.22)] \\ = 2 [0.5 - 0.4868] = 2 (0.0132) \\ = 0.0264 .$$

مثال (٩-٢) : إذا كان الزمن اللازم لجفاف طلاء من نوع ما يتبع توزيعاً طبيعياً بمتوسط 75 دليقه وانحراف معياري $\sigma = 9$ دليقة. فإذا أضيفت تعديلات على هذا الطلاء وذلك لتقليل زمن الجفاف وإذا كان الطلاء بعد التعديل لا يزال يتبع توزيعاً طبيعياً بنفس الانحراف المعياري ($\sigma = 9$) وإذا كان μ تمثل متوسط زمن الجفاف للطلاء بعد التعديل .

المطلوب : (أ) حساب احتمال الوقوع في خطأ من النوع الأول إذا كان $H_0 : \mu = 75$ و $H_1 : \mu < 75$ وذلك بالاعتماد على عينة عشوائية حجمها $n = 25$ ومنطقة الرفض $\bar{X} \leq 70.8$.

(ب) حساب احتمال الوقوع في خطأ من النوع الثاني تحت فرض العدم H_0 :

$H_1 : \mu = 70$ و $75 =$ ضد الفرض البديل أن 72

الحل . (أ)

$$\alpha = P(\text{الخطأ من النوع الأول}) \\ = P(\text{رفض } H_0 \text{ وهو صحيح}) \\ = P(\bar{X} \leq 70.8 \mid \mu = 75)$$

قيمة z المقابلة لقيمة $\bar{x} = 70.8$ في هذه الحالة هي :

$$z = \frac{70.8 - 75}{9/\sqrt{25}} = -2.33,$$

وعلى ذلك يكون :

$$\alpha = P(Z < -2.33) = P(Z > 2.33) \\ = 0.5 - P(0 < Z < 2.33)$$

$$= 0.5 - 0.4901 = 0.0099.$$

(ب) حساب β عندما $\mu = 72$

$\beta = P$ (الخطأ من النوع الثاني)

$= P$ (قبول فرض العدم وهو خطأ)

$$= P (\bar{X} > 70.8 \mid \mu = 72)$$

قيمة z المقابلة لقيمة $\bar{x} = 70.8$ في هذه الحالة هي :

$$z = \frac{70.8 - 72}{9/\sqrt{25}} = -0.67,$$

وعلى ذلك تكون :

$$\beta = P (Z > -0.67)$$

$$= 0.5 + P 0 < Z < 0.67)$$

$$= 0.5 + 0.2486 = 0.7486.$$

حساب β عند $\mu = 70$

$\beta = P$ (الخطأ من النوع الثاني)

$= P$ (قبول فرض العدم وهو خطأ)

$$= P (\bar{X} > 70.8 \mid \mu = 70)$$

قيمة z المقابلة لقيمة $\bar{x} = 70.8$ في هذه الحالة هي :

$$z = \frac{70.8 - 70}{9/\sqrt{25}} = 0.44,$$

وعلى ذلك تكون :

$$\beta = P (Z > 0.44)$$

$$= 0.5 - P 0 < Z < 0.44)$$

$$= 0.5 - 0.17 = 0.33.$$

(٣-٩) اختبارات من جانب واحد أو من جانبيين :

One - tailed and Two - tailed Tests

يسمى الاختبار ، لأى فرض إحصائي ، أنه من جانب واحد إذا كان على الصورة :

$$H_0 : \theta = \theta_0 ,$$

$$H_1 : \theta > \theta_0$$

أو على الشكل :

$$H_0 : \theta = \theta_0 ,$$

$$H_1 : \theta < \theta_0$$

منطقة الرفض للبدل $\theta > \theta_0$ تقع في الجانب الأيمن من التوزيع بينما منطقة الرفض

للبدل $\theta < \theta_0$ تقع في الجانب الأيسر من التوزيع.

يسمى الاختبار، لأي فرض إحصائي، أنه من جانبيين إذا كان على الصورة :

$$H_0 : \theta = \theta_0 ,$$

$$H_1 : \theta \neq \theta_0$$

منطقة الرفض للبدل $\theta_0 \neq \theta$ سوف تكون $\theta > \theta_0$ أو $\theta < \theta_0$. المفاضلة بين اختبار من جانب واحد أو من جانبيين سوف يتوقف على الاستنتاج الذي يرغب الباحث في الوصول إليه عند رفض فرض العدم :

في البنود التالية من هذا الفصل سوف نناقش بعض اختبارات الفروض الشائعة الاستخدام.

(٩-٤) اختبارات حول متوسط المجتمع μ

Tests About a Population Mean μ

الحالة الأولى : اختبار الفرض أن المتوسط لمجتمع يتجان معلوم σ^2 ، يساوي قيمة معينة μ_0 ضد الفرض البديل ذي جانبيين أن المتوسط لا يساوي μ_0 يكون على الشكل:

$$H_0 : \mu = \mu_0 ,$$

$$H_1 : \mu \neq \mu_0$$

الإحصاء المناسب والذي يعتمد عليه قرارنا هو المتغير العشوائي \bar{X} . إذا كان المجتمع الذي اختبرت منه العينة يتبع التوزيع الطبيعي فإن التوزيع العيني للإحصاء \bar{X} يتبع التوزيع الطبيعي

بمتوسط $\mu_{\bar{X}} = \mu$ وتباين $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ حيث أن μ و σ^2 هما المتوسط والتباين على التوالي للمجتمع الذي اختبرت منه العينات من الحجم n . عندما لا يتحقق هذا الفرض فقد

علمنا من الفصل السابع أن التوزيع العيني للإحصاء \bar{X} تقريباً يتبع التوزيع الطبيعي بمتوسط

$\mu_{\bar{X}} = \mu$ وتباين $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$. باستخدام القيمتين الحرجتين \bar{x}_1 , \bar{x}_2 فإن

$\bar{x}_1 < \bar{X} < \bar{x}_2$ تمثل منطقة القبول. الذيلين للتوزيع $\bar{X} > \bar{x}_2$ أو $\bar{X} < \bar{x}_1$ يمثلان منطقة الرفض .

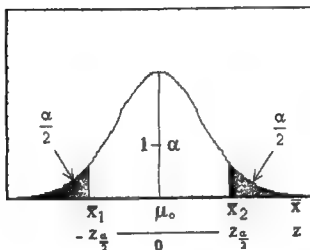
منطقة الرفض يمكن أن تغطي في صورة قيم z وذلك بعمل التحويلة التالية :

$$z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} .$$

حيث ، لمستوى معنوية α ، القيمتين الحرجتين للمطير العشوائي Z المقابلة لكل من \bar{X}_1 ، \bar{X}_2 والموضحين في شكل (٢-٩) هما :

$$-z_{\alpha/2} = \frac{\bar{X}_1 - \mu_0}{\frac{\sigma}{\sqrt{n}}},$$

$$z_{\alpha/2} = \frac{\bar{X}_2 - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$



شكل (٢-٩)

الآن لإجراء الاختبار نختار عينة عشوائية من الحجم n من المجتمع موضع الدراسة

ونحسب منها متوسط العينة \bar{X} . إذا وقعت \bar{X} في منطقة القبول $\bar{X}_1 < \bar{X} < \bar{X}_2$ فإن قيمة الإحصاء :

$$z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

تقع في المنطقة $-z_{\alpha/2} < Z < z_{\alpha/2}$. الاستنتاج سوف يكون $\mu = \mu_0$ وغير ذلك نرفض H_0

ونقبل الفرض البديل $\mu \neq \mu_0$. عادة تصاغ منطقة الرفض في صورة Z أكثر من \bar{X} . إذا كان تباین المجتمع مجهول فإننا نحسب تباین العينة s^2 ونستعمله بدلاً من σ^2 تحت شرط أن حجم العينة أكبر من أو يساوي 30 ($n \geq 30$) .

مثال (٣-٩) ينتج مصنع للأغذية المعلبة نوعا من المعلبات. قام المستولن خلال فترة طويلة بمراقبة أوزان هذه المعلبات ووجد أنها تخضع للتوزيع الطبيعي باخلاف معياري 2.6 جرام. جرت العادة في المصنع أن يكتب على العلبة الوزن الصافي وهو 300 جرام. اختيرت عينة عشوائية من 20 علبة وكان متوسط الوزن من العينة $\bar{x} = 305$. أختبر فرض العدم $\mu = 300$ ضد الفرض البديل $\mu \neq 300$ وذلك عند مستوى معنوية $\alpha = 0.01$.
الحل .

$$H_0 : \mu = 300,$$

$$H_1 : \mu \neq 300.$$

$$\alpha = 0.01 .$$

$z_{0.005} = 2.575$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣)

منطقة الرفض $Z > 2.575$ أو $Z < -2.575$

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{305 - 300}{\frac{2.6}{\sqrt{20}}} = 8.6.$$

نرفض فرض العدم لأن z تقع في منطقة الرفض.

عند الاهتمام باختبار الفرض أن متوسط المجتمع بتباين معلوم ، σ^2 ، يساوى قيمة معينة μ_0 ضد الفرض البديل من جانب واحد أي $\mu > \mu_0$ فإن فرض العدم والفرض البديل سوف يكونان على الشكل :

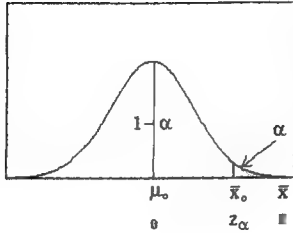
$$H_0 : \mu = \mu_0,$$

$$H_1 : \mu > \mu_0.$$

منطقة الرفض سوف تقع في الجانب الأيمن من توزيع \bar{X} . لمستوى معنوية α نحسب قيمة حرجة واحدة \bar{x}_0 بحيث $\bar{X} < \bar{x}_0$ تمثل منطقة القبول بينما $\bar{X} > \bar{x}_0$ تمثل منطقة الرفض. في شكل (٣-٩). القيمة الحرجة z_α التي تقابل القيمة \bar{x}_0 هي :

$$z_\alpha = \frac{\bar{x}_0 - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

وعلى ذلك منطقة الرفض من الحجم α تصبح $Z > z_\alpha$ للفرض البديل $\mu < \mu_0$ سوف تتبع خطوات الطريقة السابقة فيما عدا أن منطقة الرفض سوف تكون $Z < -z_\alpha$.



شكل (٣-٩)

مثال (٩-٤) من المعروف أن أحد أدوية إزالة الألم المستخدمة يمكنها إزالة الألم للمريض في فترة زمنية متوسطة 3.7 دقيقة . ولقارنة هذا الدواء بدواء جديد لإزالة الألم اختبرت عينة عشوائية من 60 مريضاً وتم إعطاء الدواء الجديد لهم فكان المتوسط الحسابي لطول فترة إزالة الألم في هذه العينة 2.2 دقيقة بانحراف معياري 1.2 دقيقة . فهل تدل هذه النتائج أن الدواء الجديد أفضل من الدواء القديم من حيث الفترة اللازمة لإزالة المرض ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.
الحل .

$$H_0 : \mu = 3.7 ,$$

$$H_1 : \mu < 3.7 .$$

$$\alpha = 0.05 .$$

وبما أن $n \geq 30$ فإنه يمكننا استخدام s بدلا من σ في صيغة z .

أن $z_{0.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣)

منطقة الرفض $Z < -1.645$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{2.2 - 3.7}{\frac{1.2}{\sqrt{60}}} = -9.682.$$

نرفض فرض العدم لأن z تقع في منطقة الرفض . أى أن الدواء الجديد أفضل من الدواء القديم من حيث الفترة اللازمة لإزالة المرض.

مثال (٩-٥) اتفق أحد مصدري البيض مع أحد التجار على أن يورد الأخير للأول عدد من البيض من الحجم الكبير ولما أحضر التاجر البيض قام المصدر باختيار عينة عشوائية مسن 100 بيضة فوجد أن متوسط وزن البيضة 67 جراما بانحراف معياري 1.6. اختبر فرض العدم $H_0: \mu = 65$ ضد الفرض البديل $H_1: \mu > 65$ (مستوى المعنوية $\alpha = 0.05$)
الحل .

$$H_0: \mu = 65,$$

$$H_1: \mu > 65.$$

$$\alpha = 0.05.$$

بما أن $n > 30$ فإنه يمكننا استخدام s بدلا من σ في صيغة z .

الحل .

$z_{0.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) .

منطقة الرفض $Z > 1.645$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{67 - 65}{\frac{1.6}{\sqrt{100}}} = 12.5.$$

نرفض H_0 لأن z تقع في منطقة الرفض. أي أن المورد سوف يقبل استلام البيض .

الحالة الثانية : في بعض اختبارات الفروض التي تخص متوسط مجتمع طبيعي عندما يكون تباين المجتمع σ^2 مجهول وحجم العينة صغير $n < 30$. استنتاجنا، في هذه الحالة، سوف يعتمد على توزيع t . بفرض اختبار من جانب واحد على الشكل :

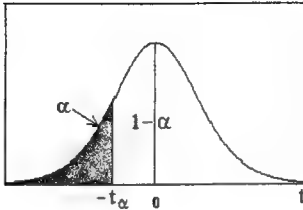
$$H_0: \mu = \mu_0,$$

$$H_1: \mu < \mu_0$$

نختار عينة عشوائية من الحجم $n < 30$ من المجتمع ونحسب المتوسط \bar{x} والانحراف المعياري s وعلى ذلك :

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

حي قيمة لتعبر عشوائي T له توزيع t بدرجات حرية $v = n - 1$ عندما H_0 صحيحاً. منطقة الرفض سوف تكون في الذيل الأيسر من توزيع t . لمستوى معنوية α يمكن الحصول على قيمة واحدة $t_{1-\alpha} = -t_\alpha$ بحيث أن $T < -t_\alpha$ تمثل منطقة الرفض و $T > -t_\alpha$ تمثل منطقة القبول. حجم منطقة الرفض يساوي المساحة المظلمة في شكل (٩-٤) .



شكل (٤-٩)

منطقة الرفض للفرض البديل $H_0: \mu > \mu_0$ بمستوى معنوية α ، هي $T > t_{\alpha}$ للفرض البديل $H_0: \mu \neq \mu_0$ فإن منطقة الرفض المقابلة لمستوى معنوية α ، هي $T > t_{\alpha/2}$ أو $T < -t_{\alpha/2}$. وعلى ذلك نحسب قيمة الإحصاء، أي قيمة t ، وإذا وقعت قيمة t في منطقة القبول نقبل فرض العدم وغير ذلك نرفض H_0 .

مثال (٦-٩) لمعرفة الرغذاء معين على زيادة الوزن اختبرت عينة عشوائية من ستة فنون وتم تغذيتها بهذا الغذاء وكانت الزيادة في أوزانهم بعد التغذية هي :

2.3 , 2.5 , 1.4 , 1.4 , 1.7 , 2.5

فهل يمكن الحكم على أن هذه العينة من مجتمع متوسط الزيادة في الوزن فيه 1.2 أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$. وذلك تحت فرض أن العينة تم اختيارها من مجتمع طبيعي .

الحل .

$$H_0: \mu = 1.2 ,$$

$$H_0: \mu \neq 1.2 .$$

$$\alpha = 0.05 .$$

$t_{0.025} = 2.571$ والمستخرجة من توزيع t في ملحق (٤) عند درجات حرية $v = 5$.

منطقة الرفض $T > 2.571$ أو $T < -2.571$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{11.8}{6} = 1.967 ,$$

$$s = \sqrt{\frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n} \right]}$$

$$= \sqrt{\frac{1}{5} \left[24.6 - \frac{(11.8)^2}{6} \right]} = 0.528.$$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{1.967 - 1.2}{\frac{0.528}{\sqrt{6}}} = 3.5582.$$

نرفض H_0 لأن t تقع في منطقة الرفض .

(٩-٥) اختبارات حول تباين المجتمع σ^2

Tests about the Population Variance σ^2

عند الرغبة في اختبار الفرض أن التباين مجتمع طبيعي يساوى قيمة معينة σ_0^2 ضد الفرض البديل ذي جانبيين أن التباين لا يساوى σ_0^2 . أي أننا نختبر الفرض أن :

$$H_0 : \sigma^2 = \sigma_0^2 ,$$

$$H_1 : \sigma^2 \neq \sigma_0^2$$

نختار عينة عشوائية من الحجم n من المجتمع موضع الدراسة ونحسب تباين العينة s^2 . وعلى ذلك:

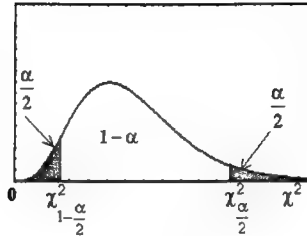
$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

قيمة للمتغير χ^2 ، بما لنظرية (٧ - ١٢) ، والذي له توزيع χ^2 بدرجات حرية $v = n - 1$

عندما يكون H_0 صحيحاً . مستوى معنوية α نوجد القيمتين الحرجتين $\chi_{1-\frac{\alpha}{2}}^2$ ، $\chi_{\frac{\alpha}{2}}^2$ بحيث أن

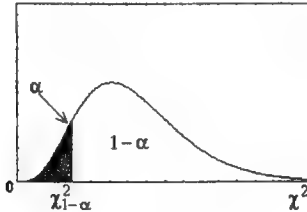
$\chi^2 < \chi_{1-\frac{\alpha}{2}}^2$ ، $\chi^2 > \chi_{\frac{\alpha}{2}}^2$ يمثلان منطقة الرفض . حجم المنطقة الحرجة يساوى المساحة المظلمة

في شكل (٩-٥) . نرفض H_0 إذا وقعت χ^2 في منطقة الرفض .



شكل (٥-٩)

عادة يكون الاهتمام باختيار الفرض $\sigma^2 = \sigma_0^2$ ضد الفرض البديل من جانب واحد من التوزيع . للبديل $\sigma^2 < \sigma_0^2$ والمستوى معنوية α نحصل على قيمة حرجية $\chi^2_{1-\alpha}$ بحيث أن $X^2 < \chi^2_{1-\alpha}$ تمثل منطقة الرفض و $X^2 > \chi^2_{1-\alpha}$ تمثل منطقة القبول. بنفس الشكل لبديل من جانب واحد $\sigma^2 > \sigma_0^2$ ، فإن χ^2_{α} تمثل القيمة الحرجية بحيث $X^2 > \chi^2_{\alpha}$ تمثل منطقة الرفض و $X^2 < \chi^2_{\alpha}$ تمثل منطقة القبول. حجم المنطقة الحرجية لبديل من جانب واحد $\sigma^2 < \sigma_0^2$ يساوي المساحة المظللة في شكل (٦-٩) .



شكل (٦-٩)

مثال (٧-٩) يعتقد مسئول في مصنع لبطاريات السيارات أن الانحراف المعياري لعمر البطارية المنتجة هو ٠.٨ . لاختبار ذلك الفرض اختبرت عينة عشوائية من الحجم $n = 20$ وتمت تجربتها فكان $s = 1.1$ سنة فهل يمكن القول أن $\sigma > 0.8$ ؟ (استخدم مستوى معنوية $\alpha = 0.05$).
الحل .

$$H_0: \sigma^2 = 0.64,$$

$$H_1: \sigma^2 > 0.64$$

$$\alpha = 0.05.$$

$\chi^2_{0.05} = 30.143$ والمستخرجة من جدول توزيع χ^2 بدرجات حرية $v = 19$. منطقة

الرفض $X^2 > 30.143$.

$$\begin{aligned}\chi^2 &= \frac{(n-1)s^2}{\sigma_0^2} = \frac{(19)(1.1)^2}{0.8^2} \\ &= 35.922.\end{aligned}$$

بما أن χ^2 تقع في منطقة الرفض نرفض H_0 .

(٩-٦) اختبارات فحص تبايني مجتمعين

Tests Concerning Two Populations Variances

يفرض أن لدينا مجتمعين : الأول يتبع توزيعاً طبيعياً متوسطة μ_1 وتباينه σ_1^2 والثاني : يتبع توزيعاً طبيعياً متوسطة μ_2 وتباينه σ_2^2 والمطلوب اختبار هل المجتمعين لهما نفس التباين ؟ أى هل $\sigma_1^2 = \sigma_2^2$ أم لا ؟ فإذا كانت $\sigma_1^2 = \sigma_2^2$ فإننا نقول أن هناك تجانس بين المجتمعين . إن التأكد من صحة الفرض $\sigma_1^2 = \sigma_2^2$ ضروري لاختبار الفرق بين متوسطي مجتمعين (اختبار t) والذي سوف نتناوله في البند التالي . أيضا هناك العديد من الأبحاث التي يكون هدفها الرئيسي هو مقارنة σ_1^2 مع σ_2^2 مثل دراسات جودة البضائع المستهلكة حيث يعتبر التباين أهم مقياس الجودة.

لاختبار فرض العدم :

$$H_0: \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

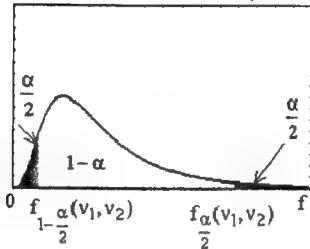
$$H_1: \sigma_1^2 \neq \sigma_2^2$$

نختار عينة عشوائية حجمها n_1 من المجتمع الأول ولكن متوسطها الحسابي \bar{x}_1 وتباينها s_1^2 ونختار عينة عشوائية أخرى حجمها n_2 من المجتمع الثاني ولكن متوسطها \bar{x}_2 وتباينها s_2^2 .
(العينة الثانية مستقلة عن العينة الأولى) . بالفرض صحة فرض العدم فإن :

$$f = \frac{s_1^2}{s_2^2}$$

تمثل قيمة للمتغير العشوائي F والذي تبعاً لنظرية (٧-١٣) له توزيع F بدرجات حرية $v_1 = n_1 - 1, v_2 = n_2 - 1$ مستوى معنوية α ، سوف نحصل على قيمتين حرجيتين $f_{\alpha}(v_1, v_2)$ و $f_{1-\frac{\alpha}{2}}(v_1, v_2)$ وعلى ذلك فإن $F > f_{\alpha}(v_1, v_2)$ أو $F < f_{1-\frac{\alpha}{2}}(v_1, v_2)$ تعادلان منطقة الرفض. حجم منطقة الرفض يساوي المساحة المظللة في شكل (٧-٩). باستخدام نظرية (٧-١٤) فإن القيمة الحرجة للمتغير F في الطرف الأيسر يمكن الحصول عليها من العلاقة التالية :

$$f_{1-\frac{\alpha}{2}}(v_1, v_2) = \frac{1}{f_{\alpha}(v_2, v_1)}.$$



شكل (٧-٩)

إذا وقعت f المحسوبة في منطقة الرفض فإننا نرفض H_0 .

لاختبار فرض العدم :

$$H_0 : \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

$$H_1 : \sigma_1^2 < \sigma_2^2$$

فإن منطقة الرفض ، بمستوى معنوية α ، سوف تكون في الجانب الأيسر من التوزيع (الذيل الأيسر) . منطقة الرفض في هذه الحالة تمثل كل قيم F بحيث $F < f_{1-\alpha}(v_1, v_2)$. وأخيرا لاختبار فرض العدم :

$$H_0 : \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

$$H_1 : \sigma_1^2 > \sigma_2^2$$

فإن منطقة الرفض ، بمستوى معنوية α ، سوف تكون في الجانب الأيمن من التوزيع (الذيل الأيمن) . منطقة الرفض في هذه الحالة تمثل كل قيم F بحيث $F > f_{\alpha}(v_1, v_2)$. مثال (٨-٩) من البيانات التالية اختبر التجانس بين المجتمعين وذلك عند مستوى معنوية

$$\alpha = 0.1$$

	العينة الثانية	العينة الأولى
s_j^2	50.7	40.5
n_j	41	31

الحل .

$$H_0 : \sigma_1^2 = \sigma_2^2 ,$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2 .$$

$$\alpha = 0.1 .$$

$f_{0.95}(40, 30) = 1.79$ والمستخرجة من جدول توزيع F في ملحق (٦) عند

درجات حرية $v_1 = 40, v_2 = 30$. أما $f_{0.95}(40, 30)$ فتحسب من العلاقة التالية :

$$f_{0.95}(40, 30) = \frac{1}{f_{0.05}(30, 40)} = \frac{1}{1.74} = 0.575 .$$

منطقة الرفض $F < 0.575$ أو $F > 1.79$

التباين الأكبر

$$f = \frac{s_2^2}{s_1^2} = \frac{50.7}{40.5} = 1.252 .$$

التباين الأصغر

نقبل H_0 لأن f تقع في منطقة القبول .

(٧-٩) اختبارات تخص المتوسطات Tests Concerning Means

في بعض الأحيان يكون الاهتمام باختبارات الفروض التي تخص مجتمعين مختلفين. أي أننا نرغب في اختبار فرض العدم أن الفرق بين متوسطي مجتمعين ، $\mu_1 - \mu_2$ ، يساوى صفر أي $\mu_1 = \mu_2$ ضد الفرض البديل $\mu_1 - \mu_2 > 0$ أي $\mu_1 > \mu_2$ أو الفرض البديل $\mu_1 - \mu_2 < 0$ أي $\mu_1 < \mu_2$ أو الفرض البديل $\mu_1 - \mu_2 \neq 0$ أي $\mu_1 \neq \mu_2$. تعتمد الطريقة المستخدمة في اختبار الفرق بين متوسطي مجتمعين على توزيع كل مجتمع وحجم العينة المختارة من كل مجتمع. في الجزء التالي سوف نتناول ثلاثة حالات.

الحالة الأولى: عند اختبار فرض العدم H_0 أن الفرق بين متوسطي مجتمعين ،

$\mu_1 - \mu_2$ ، يساوى صفر وذلك عندما كل من σ_1^2 ، σ_2^2 معلومتان وتحت فرض أن كل مجتمع له توزيعاً طبيعياً أو تقريباً طبيعياً. أما في حالة العينات الكبيرة وإذا كانت σ_1^2 ، σ_2^2 مجهولتان فإنه يمكن تقديرهما من العينات بحساب s_1^2 ، s_2^2 . يعتمد قرارنا في هذه الحالة على المتغير العشوائي (الإحصاء) $\bar{X}_1 - \bar{X}_2$. أولاً نختار عينة عشوائية من الحجم n_1 من المجتمع الأول ونحسب منها \bar{x}_1 ونختار عينة عشوائية أخرى من الحجم n_2 من المجتمع الثاني (مستقلة عن العينة الأولى) ونحسب منها \bar{x}_2 ثم نحسب الفرق ، $\bar{x}_1 - \bar{x}_2$ ، لمتوسطي العينةين ، من نظرية (٦-٧) نعلم أن :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

قيمة للمتغير العشوائي Z عندما يكون H_0 صحيحاً. وعلى ذلك في اختبار من جانبيين وعند مستوى معنوية α فإن منطقة الرفض تحدد على الشكل $Z > z_{\frac{\alpha}{2}}$ أو $Z < -z_{\frac{\alpha}{2}}$. أما في

اختبار من جانب واحد حيث الفرض البديل $\mu_1 < \mu_2$ فإن منطقة الرفض ، مستوى معنوية α ، سوف تكون $Z < -z_{\alpha}$. وأخيراً في حالة الفرض البديل من جانب واحد $\mu_1 > \mu_2$ فإن منطقة الرفض مستوى معنوية α ، سوف تكون $Z > z_{\alpha}$.

مثال (٩-٩) أجرى اختبار على المقاومة للشد tensile strength لنوعين من السلك . النتائج معطاة في الجدول التالي :

النوع	حجم العينة	متوسط العينة	الانحراف المعياري للعينة
A	$n_1 = 129$	$\bar{x}_1 = 107.6$	$s_1 = 1.3$
B	$n_2 = 129$	$\bar{x}_2 = 123.6$	$s_2 = 2.0$

المطلوب اختبار هل هناك فرقاً معنوياً بين متوسطي المجتمعين المسحوبين منهما العينتين ؟ (عند مستوى معنوية $\alpha = 0.05$).

الحل . حيث أن $n_1 > 30$ و $n_2 > 30$ نضع الآتي :

$$H_0 : \mu_1 = \mu_2 ,$$

$$H_1 : \mu_1 \neq \mu_2 .$$

$$\alpha = 0.05 .$$

$z_{0.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) .

منطقة الرفض $Z > 1.96$ أو $Z < -1.96$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$= \frac{107.6 - 123.6}{\sqrt{\frac{1.3^2}{129} + \frac{2^2}{129}}} = -76.183.$$

وبما أن z تقع في منطقة الرفض فإننا نرفض H_0 .

الحالة الثانية : بفرض أن σ_1^2 , σ_2^2 مجهولتان وحجم كلا من العينتين صغير . يعتمد

القرار الذي نتخذه في هذه الحالة في اختبار فرض العدم على توزيع t وذلك تحت فرض

$\sigma_1^2 = \sigma_2^2 = \sigma^2$ (هناك تجانس) وأن كل مجتمع له توزيعاً طبيعياً . أولاً نختار عينة

عشوائية من الحجم n_1 من المجتمع الأول ونحسب منها \bar{x}_1 , s_1^2 ونختار عينة عشوائية أخرى من

الحجم n_2 من المجتمع الثاني (مستقلة عن العينة الأولى) ونحسب منها \bar{x}_2 , s_2^2 .

التباين التجميعي **pooled variance** نحصل عليه من الصيغة التالية :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

من نظرية (٧-١٠) وتحت فرض أن H_0 صحيحاً فلنأخذ نعلم أن :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

قيمة لمغير عشوائي T يتبع توزيع t بدرجات حرية $v = n_1 + n_2 - 2$ في حالة اختبار

ذي جانبيين وعند مستوى معنوية α فإن منطقة الرفض سوف تكون $T > t_{\frac{\alpha}{2}}$ أو

$T < -t_{\frac{\alpha}{2}}$. للبدل من جانب واحد $\mu_1 < \mu_2$ فإن منطقة الرفض سوف تكون $T < -t_{\frac{\alpha}{2}}$.

وأخيراً للبدل $\mu_1 > \mu_2$ فإن منطقة الرفض سوف تكون $T > t_{\frac{\alpha}{2}}$.

مثال (٩-١٠) اختبرت مجموعتان من الطلبة وأعطيت المجموعة الأولى الوجبة A يومياً أعطيت المجموعة الثانية الوجبة B يومياً . وقد استمرت التجربة لمدة شهر وكانت الزيادة في وزن مفردات كل مجموعة (بالرطل) هي :

المجموعة الأولى 2.6, 2.7, 3.9, 3.4, 1.0, 1.6, 4.0, 3.6, 2.4, 3.0

المجموعة الثانية 2.9, 1.4, 2.6, 1.9, 1.9, 2.4, 2.9, 3.6, 1.6

فهل تعتقد أن هناك فرقاً معنوية بين تأثير الوجبة A والوجبة B على زيادة الوزن ؟ (عند مستوى معنوية $\alpha = 0.1$) ، وذلك تحت فرض أن المئين تم اختبارهما من مجتمعين طبيعيين .
الحل .

$$n_1 = 10, \bar{x}_1 = 2.820, s_1 = 0.976$$

$$n_2 = 9, \bar{x}_2 = 2.356, s_2 = 0.716$$

أولاً يجب التأكد من أن $\sigma_1^2 = \sigma_2^2$ أى اختبار فرض العلم :

$$H_0 : \sigma_1^2 = \sigma_2^2$$

ضد الفرض البديل :

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$\alpha = 0.1$$

التباين الأكبر

$$f = \frac{s_1^2}{s_2^2} = \frac{(0.976)^2}{(0.716)^2} = 1.858$$

التباين الأصغر

$f_{0.05}(9, 8) = 3.39$ والمستخرجة من جدول توزيع F في ملحق (٦) عند درجات

حرية $v_1 = 9, v_2 = 8$ أما $f_{0.95}(9, 8)$ فتحسب من العلاقة التالية :

$$f_{0.95}(9, 8) = \frac{1}{f_{0.05}(8, 9)} = \frac{1}{3.23} = 0.3096.$$

منطقة الرفض $F < 0.3096$ أو $F > 3.39$

وبما أن f تقع في منطقة القبول فإننا نقبل فرض العدم أن $\sigma_1^2 = \sigma_2^2$.

الآن نختصر :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

$$\alpha = 0.1.$$

$$s_p = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{(0.976)^2(9) + (0.716)^2(8)}{10 + 9 - 2}}$$

$$= \sqrt{0.7455548} = 0.86346.$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{2.820 - 2.356}{0.86346 \sqrt{\frac{1}{10} + \frac{1}{9}}} = 1.16955.$$

$t_{0.05} = 1.74$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $v = 17$.

منطقة الرفض $T > 1.74$ أو $T < -1.74$. وبما أن t تقع في منطقة القبول فإننا نقبل H_0

وهذا يدل على عدم وجود فرق معنوي بين تأثير الوجبة A والوجبة B على زيادة الوزن .

الحالة الثالثة : عند اختبار فرض العدم $H_0 : \mu_1 = \mu_2$ ضد الفرض البديل $H_1 : \mu_1 \neq \mu_2$

تحت الشروط التالية :

(أ) كل مجتمع (من المجتمعين تحت الدراسة) يتبع توزيعاً طبيعياً .

(ب) تباين المجتمعين ، $\sigma_1^2 = \sigma_2^2$ ، مختلفين كثيراً .

(ج) العينتان صغيرتان وإحجامهما مختلفتان .

القرار الذي نتخذه يعتمد على الإحصاء T' والذي تقريباً يتبع توزيع t بدرجات حرية محسوب من الصيغة التالية :

$$v = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]^2}{\left[\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 - 1} \right]}$$

لإجراء الاختبار نختار عينة عشوائية حجمها n_1 من المجتمع الأول كما نختار عينة عشوائية أخرى حجمها n_2 من المجتمع الثاني (العينة الثانية مستقلة عن العينة الأولى). نحسب $\bar{x}_1, \bar{x}_2, s_1^2, s_2^2$ وعلى ذلك يكون :

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

قيمة للمتغير T' عندما يكون H_0 صحيحاً. لاختبار ذي جانبيين ، بمستوى معنوية α ، منطقة الرفض تقريباً تعطى حيث $t_{\frac{\alpha}{2}}$ أو $-t_{\frac{\alpha}{2}}$ هما القيمتين الخارجيتين لتوزيع t بدرجات حرية v منطقة الرفض سوف تكون $T' > t_{\frac{\alpha}{2}}$ أو $T' < -t_{\frac{\alpha}{2}}$. لبدل من جانب واحد $\mu_1 < \mu_2$ ، فإن منطقة الرفض سوف تكون $T' < -t_{\alpha}$ وللبديل $\mu_1 > \mu_2$ فإن منطقة الرفض سوف تكون $T' > t_{\alpha}$.

مثال (٩-١١) أوضحت الدراسة أن زيادة النترات nitrate في الاستهلاك الآدمي له تأثيرات ضارة منها قلة إنتاج الثيروكسين وقلة إدرار اللبن عند البقر. البيانات التالية نتيجة تجربة لقياس النسبة المئوية للزيادة في وزن فئران تجارب صغيرة العمر تناولت وجبة قياسية وفئران تناولت 2000 ppm نترات من مياه الشرب.

17.2	14.0	16.6	10.8	14.0	10.5	20.5	19.3	12.7	النترات
15.7	27.6	16.2	14.3	10.0	32.9	18.2	المراقبة (القياسية)		

تحقق من صحة الفرض القائل : لا يوجد فرق معنوي بين مجموعة التواتر ومجموعة المراقبة وذلك عند مستوى معنوية $\alpha = 0.1$. (تحت فرض أن العيتين تم اختيارهما من مجتمعين طبيعيين) .

الحل . يجب علينا أولاً التحقق من $\sigma_1^2 = \sigma_2^2$.

$s_1 = 3.558$, $s_2 = 8.053$ وعلى ذلك فإن قيمة f هي :

التباين الأكبر

$$f = \frac{s_2^2}{s_1^2} = \frac{(8.053)^2}{(3.558)^2} = 5.1228.$$

التباين الأصغر

$f_{0.05}(6, 8) = 3.58$ والمستخرجة من جدول توزيع F في ملحق (٦) عند درجات

حرية $v_1 = 6, v_2 = 8$. أما $f_{0.95}(6, 8)$ فيمكن الحصول عليها من العلاقة التالية :

$$f_{0.95}(6,8) = \frac{1}{f_{0.05}(8,6)} = \frac{1}{4.15} = 0.241.$$

منطقة الرفض $F < 0.241$ أو $F > 3.58$

وحيث أن f تقع في منطقة الرفض فإننا نرفض H_0 ونستنتج أن $\sigma_1^2 \neq \sigma_2^2$ الآن نختصر :

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

$$\alpha = 0.1.$$

$$\bar{x}_2 = 19.271 \quad , \quad \bar{x}_1 = 15.067$$

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

$$= \frac{15.067 - 19.271}{\sqrt{\frac{3.558^2}{9} + \frac{8.053^2}{7}}} = -1.2869.$$

وعلياً أن نقارن قيمة t' المحسوبة بقيمة t الجدولية عند درجات حرية :

$$v = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]^2}{\left[\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 - 1} \right]}$$

$$= \frac{\left[\frac{3.558^2}{9} + \frac{8.053^2}{7} \right]^2}{\left[\frac{\left(\frac{3.558^2}{9} \right)^2}{8} + \frac{\left(\frac{8.053^2}{7} \right)^2}{6} \right]}$$

$$= \frac{113.87}{14.55} = 7.8 \approx 8.$$

$t_{0.05} = 1.86$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $v = 8$ وعلى ذلك فإن منطقة الرفض $T' > 1.86$ أو $T' < -1.86$. بما أن t تقع في منطقة القبول نقبل H_0 وهذا يعني عدم وجود فرق معنوي بين مجموعة الثمرات ومجموعة المراقبة عند مستوى معنوية $\alpha = 0.1$.

(٨-٩) اختبارات t للأزواج The Paired t Tests

في البند (٩-٧) كان اهتمامنا بالعينات المستقلة . الآن سوف يكون اهتمامنا بالعينات المزدوجة **paired samples** ، حيث d_1, d_2, \dots, d_n تمثل الفروق لأزواج المشاهدات المرتبطة التي عددها n . مثل هذه المشاهدات تحدث عندما نأخذ المشاهدات (القراءات) على الأفراد قبل وبعد معالجة . بالنظر إلى الفروق لكل أزواج المشاهدات فإننا نأمل في الوصول إلى استنتاج يخص تأثير المعالجة . متوسط فروق المجتمع ، μ_D ، سوف يساوي الفرق بين متوسطي المجتمعين ، $\mu_1 - \mu_2$ ، وعلى ذلك فإن مشكلة اختبار فرض العدم H_0 أن $\mu_1 - \mu_2 = 0$ تكافئ اختبار $\mu_D = 0$. بفرض أن كل مجتمع من المجتمعين يتبع توزيعاً طبيعياً . أولاً نختار n

من أزواج المشاهدات عشوائيا ونحسب الفروق ونقدر \bar{d} و s_d . وعلى ذلك تبعاً لنظرية (٧-١١) ، فإننا نعلم أنه عندما يكون H_0 صحيحاً فإن :

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}}$$

هي قيمة للمتغير العشوائي T الذي يتبع توزيع t بدرجات حرية $v = n - 1$. قرارنا في هذه الحالة يعتمد على توزيع t ومناطق الرفض المقابلة للفروض البديلة المختلفة وسوف نتبع الخطوات التي أتبعنا من قبل .

مثال (٩-١٢) إذا كانت أوزان 10 أشخاص قبل التوقف عن التدخين وبعد 8 أسابيع من التوقف عن التدخين كما يلي :

الأشخاص	1	2	3	4	5	6	7	8	9	10
قبل	148	176	152	115	166	147	150	175	151	146
بعد	155	170	169	117	170	150	154	180	160	151

فهل تدل هذه البيانات على أن الامتناع عن التدخين يؤدي إلى زيادة وزن الأشخاص الذين يمتنعون عن التدخين ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.
الحل .

$$H_0 : \mu_D = 0,$$

$$H_1 : \mu_D \neq 0.$$

$$\alpha = 0.05.$$

$t_{0.025} = 2.262$ والمستخرجة من جدول توزيع t في ملحق (٤) عند درجات حرية $v = 9$.

منطقة الرفض $T > 2.262$ أو $T < -2.262$

$$\bar{d} = \frac{\sum_{i=1}^n x_i}{n} = \frac{-50}{10} = -5 ,$$

$$s_d = \sqrt{\frac{1}{n-1} \left[\sum_{i=1}^n d_i^2 - \frac{(\sum_{i=1}^n d_i)^2}{n} \right]}$$

$$= \sqrt{\frac{1}{9} \left[550 - \frac{(-50)^2}{10} \right]} = 5.7735,$$

$$t = \frac{\bar{d} - 0}{\frac{s_d}{\sqrt{n}}} = \frac{-5}{5.7735/\sqrt{10}} = -2.7386.$$

وبما أن t تقع في منطقة الرفض نرفض H_0

(٩-٩) اختبارات تخص نسبة مجتمع

Tests Concerning a Population Proportion

سوف نتم في هذا البند بمشكلة اختبارات الفروض التي فيها نسبة صفة ما تساوي قيمة

معينة . أى أننا نتم باختبار فرض العدم $H_0 : p = p_0$ ضد الفرض البديل $p < p_0$ أو $p > p_0$. سوف تقتصر دراستنا في هذا البند على حالة العينات الكبيرة ، وعلى ذلك فإن الإحصاء المناسب الذي يعتمد عليه قرارنا هو \hat{P} الذي تقريباً يتبع توزيعاً طبيعياً . أى أن قرارنا سوف يعتمد على :

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

والذي يمثل قيمة للمتغير Z الذي يتبع التوزيع الطبيعي القياسي . وعلى ذلك للاختبار من جانبين فإن منطقة الرفض ، بمستوى معنوية α ، سوف تكون $Z > z_{\frac{\alpha}{2}}$ أو $Z < -z_{\frac{\alpha}{2}}$.
للبديل من جانب واحد $p < p_0$ فإن منطقة الرفض سوف تكون $Z < -z_{\alpha}$. وأخيراً للبديل $p > p_0$ فإن منطقة الرفض سوف تكون $Z > z_{\alpha}$.

مثال (٩ - ١٣) يدعى منتج أن ٩٠% من قطع الغيار التي يمد بها مصنعاً مطابقاً للمواصفات . فإذا تم اختبار عينة عشوائية من ٢٠٠ قطعة ووجد أن ٤٠ قطعة تالفة . أختبر ادعاء المنتج عند مستوى معنوية $\alpha = 0.05$.

$$H_0 : p = 0.9,$$

$$H_1 : p \neq 0.9.$$

$$\alpha = 0.05.$$

$z_{0.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣). منطقة

الرفض $Z > 1.96$ أو $Z < -1.96$

$$x = 160 , \quad n = 200.$$

$$\hat{p} = \frac{x}{n} = \frac{160}{200} = 0.8, \quad q_0 = 1 - q_0 = 1 - 0.9 = 0.1.$$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.8 - 0.9}{\sqrt{\frac{(0.9)(0.1)}{200}}} = -4.714.$$

وبما أن z تقع في منطقة الرفض نرفض H_0 .

(٩-١٠) اختبارات تخص الفرق بين نسبي مجتمعين

Tests Concerning a Difference Between Two Population Proportions

بفرض أن p_1 هي نسبة توفّر صفة ما في إحدى المجتمعات وكانت p_2 هي نسبة توفّر الصفة نفسها في مجتمع آخر وإذا كان اهتمامنا باختبار لفرض العدم $H_0: p_1 = p_2$ فإن الإحصاء المناسب والذي يعتمد عليه قرارنا سوف يكون المتغير العشوائي $\hat{P}_1 - \hat{P}_2$. نختار عينة عشوائية كبيرة من الحجم n_1 من المجتمع الأول ونحسب نسبة توفّر الصفة محل الدراسة فيها ولكن $\hat{p}_1 = \frac{x_1}{n_1}$ حيث أن x_1 هي عدد الذين يحملون الصفة في المجتمع الأول. نختار عينة عشوائية أخرى كبيرة من الحجم n_2 من المجتمع الثاني ونحسب نسبة توفّر الصفة المطلوبة منها ولكن $\hat{p}_2 = \frac{x_2}{n_2}$ حيث x_2 هي عدد الذين يمتلكون الصفة في المجتمع الثاني، ويجب أن تكون العنيتين مستقلتين تحت فرض العدم ومن نظرية (٧-٨) فإن :

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}.$$

هو قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي عندما H_0 يكون صحيحا و n_1, n_2 كبيرتان. وبما أن p مجهولة في صيغة z فإننا نحسبها من الصيغة التالية :

$$\tilde{p} = \frac{x_1 + x_2}{n_1 + n_2}.$$

وعلى ذلك تصبح z كالتالي :

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\tilde{p} \tilde{q} \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

حيث أن $\tilde{q} = 1 - \tilde{p}$

منطقة الرفض للفروض البديلة المختلفة يمكن الحصول عليها ، كما سبق أن ذكرنا ، باستخدام القيم الحرجة لمنحنى التوزيع الطبيعي القياسي .

مثال (٩ - ١٤) اختبرت عينة عشوائية من 300 مدخنا في مدينة ما ووجد أن من بينهم 60 يفضلون تدخين النوع A من السجائر ثم اختبرت عينة عشوائية من 200 مدخنا في مدينة أخرى ووجد أن من بينهم 30 يفضلون تدخين النوع A من السجائر . اختبر فرض العدم $H_0 : p_1 = p_2$ ضد الفرض البديل $H_1 : p_1 \neq p_2$ وذلك عند مستوى معنوية $\alpha = 0.05$.

$$H_0 : p_1 = p_2 , \quad \text{الحل} .$$

$$H_1 : p_1 \neq p_2 .$$

$$\alpha = 0.05 .$$

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{60}{300} = 0.2 , \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{30}{200} = 0.15 ,$$

$$\tilde{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{60 + 30}{300 + 200} = \frac{90}{500} = 0.18 , \quad \tilde{q} = 1 - \tilde{p} = 1 - 0.18 = 0.82 .$$

$z_{0.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) .

منطقة الرفض $Z > 1.96$ أو $Z < -1.96$

$$z = \frac{(0.2 - 0.15) - 0}{\sqrt{(0.18)(0.82) \left[\left(\frac{1}{300} \right) + \left(\frac{1}{200} \right) \right]}} = 1.4257 .$$

بما أن z تقع في منطقة القبول فإننا نقبل H_0 .

تأارين :

١- بين أي الجمل التالية صواب وأيهم خطأ .

$$A - H_0 : S \leq 0.2 \quad \text{ب} - H_0 : \bar{x} = 45$$

$$ج - H_0 : \mu = 100 , H_1 : \mu > 100$$

$$H_0 : \bar{X}_1 - \bar{X}_2 = 5 \quad \text{د -}$$

$$H_0 : p = 0.25 , H_1 : p \neq 0.25 \quad \text{هـ -}$$

$$H_0 : S_1^2 = S_2^2 , H_1 : S_1^2 \neq S_2^2 \quad \text{و -}$$

٢- إذا كانت μ تمثل المتوسط الحقيقي لمسوى الإشعاع (مقاسه Picocuries per liter

(. القيمة 5 pci/L تمثل القيمة الحرجة (الحد الفاصل بين الأمان وعدم الأمان للماء) هل تقترح

فرض العدم $H_0 : \mu = 5$ ضد الفرض البديل $H_1 : \mu > 5$ أو اختبار فرض العدم

$H_0 : \mu = 5$ ضد الفرض البديل $H_1 : \mu < 5$ ولماذا ؟

٣- إذا كان الإحصاء Z يتبع توزيع طبيعي قياسي عندما تكون H_0 صحيح أوجد مستوى

المعنوية لكل من الحالات التالية :

$$\text{أ- } H_1 : \mu > \mu_0 \quad \text{ومنطقة الرفض } Z > 1.88$$

$$\text{ب- } H_1 : \mu < \mu_0 \quad \text{ومنطقة الرفض } Z < -2.75$$

$$\text{ج- } H_1 : \mu \neq \mu_0 \quad \text{ومنطقة الرفض } Z > 2.88 \text{ أو } Z < -2.88$$

٤- مجتمع طبيعي له متوسط μ وانحراف معياري $\sigma = 0.5$. لاختبار فرض العدم

$H_0 : \mu = 15$ اخترت عينة عشوائية من الحجم $n = 25$ وكانت منطقة الرفض $\bar{X} < 14.9$

أوجد :

أ- احتمال الوقوع في خطأ من النوع الأول ؟

ب- احتمال الوقوع في خطأ من النوع الثاني إذا كان الفرض البديل $\mu = 14.8$ و

$\mu = 14.9$.

٥- تص المواصفات القياسية على أن قوة الضغط Compression Strength خليط

من الأسمنت البورتلاندى ومركبات أخرى لا بد أن تزيد عن 1300 k N/m . بفرض أن قوة

الضغط لهذا الخليط يتبع توزيعاً طبيعياً بمتوسط μ وانحراف معياري $\sigma = 60$.

(أ) ما هو فرض العدم والفرض البديل المناسب لاختبار μ في هذه الحالة ؟

(ب) إذا كانت \bar{X} تمثل الإحصاء المناسب واخترت عينة عشوائية من الحجم $n = 25$ وإذا

كانت منطقة الرفض $\bar{X} \geq 1331.26$. ما هو احتمال الوقوع في خطأ من النوع الأول ؟

٦- إذا كانت النسبة المئوية المرغوبة لمركب SiO_2 في نوع معين من الأسمنت (في المتوسط

(هي 5.5 . لاختبار فرض العدم $H_0 : \mu = 5.5$ اخترت 16 عينة مستقلة للتحليل . بفرض

أن النسبة المئوية للمركب SiO_2 في العينة تتبع توزيعاً طبيعياً بانحراف معياري $\sigma = 0.3$

وإذا كانت $\bar{x} = 5.25$ اختبر فرض العدم $H_0: \mu = 5.5$ ضد الفرض البديل $H_1: \mu \neq 5.5$ عند مستوى معنوي $\alpha = 0.01$.

٧- يتطلب العمليات الجراحية التي تجري للحيول في ظروف الحقل مخدر يؤدي إلى تخدير الحيوان لزم من مخدر. اختبرت عينة عشوائية من 73 حيوانا كانوا تحت المخدر وتم حساب أزمنة التخدير لهم وكان $\bar{x} = 18.86$ دقيقة بالمخدرات معياري $s = 8.6$ دقيقة هل هذه البيانات تدل على أن متوسط زمن الغياب عن الوعي تحت ظروف الحقل أقل من 20 دقيقة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٨- يوصي معهد التغذية على أن كمية الزنك التي يحتاج إليها الفرد الذكر في العمر أكبر من 50 سنة هو 15 mg/day أجريت تجربة وتم الحصول على النتائج التالية :

$$\bar{x} = 11.3, \quad n = 115, \quad s = 6.45$$

هل تدل هذه البيانات على أن متوسط الاحتياج اليومي من الزنك في مجتمع الذكور من العمر أكبر من 50 تقل عن المسموح به ؟

٩- تقوم شركة لصناعة الإطارات بإنتاج نوع من إطارات السيارات التي يتحمل ضغط 30 lb/in^2 . فإذا كانت μ هو المتوسط الحقيقي للضغط. أوجد مستوى المعنوية α لتقييم z التالية :

$$\text{أ- } 2.10 \quad \text{ب- } 1.75 \quad \text{ج- } 1.41 \quad \text{د- } 5.3$$

١٠- إذا كانت μ تمثل متوسط سير الدم لكل النساء الحوامل. لاختبار فرض العدم $H_0: \mu = 5.63$ ضد الفرض البديل $H_1: \mu \neq 5.63$ وذلك بالاعتماد على عينة عشوائية من النساء الحوامل من الحجم $n = 176$. هل تعتقد أنه يمكن رفض H_0 عند مستوى معنوية $\alpha = 0.01$ إذا كانت $\bar{x} = 6.1$, $s = 0.7$ ؟

١١- تقوم شركة لتصنيع الأدوية بإنتاج نوع معين من الأسيرين في عبوات سعة العبوة 100 حبة. قسم الشركة بوزن الحبة أكثر من اهتمامها بالعدد، فإذا كان من الضروري أن يكون متوسط وزن الحبة 5 grains. اختبرت عينة عشوائية من الحجم $n = 100$ حبة من إنتاج كبير وكان متوسط وزن الحبة $\bar{x} = 4.87$ بالمخدرات معياري $s = 0.35$. هل المعلومات التي تم الحصول عليها كافية لاستنتاج أن الشركة تنتج طبقاً للمواصفات القياسية ؟ اختبر الفرض المناسب عند مستوى معنوية 0.01.

١٢- أوضحت الخبرة الماضية أن درجات الطلبة في مادة الإحصاء تتبع توزيعاً طبيعياً بمتوسط 75 وتباين 16. يرغب الأعضاء في قسم الرياضيات في معرفة هل متوسط درجات الطلبة في

السنة الحالية لها نفس مستوى السنوات الماضية ؟ لذلك قررروا اختبار فرض المدم $H_0: \mu = 75$ ضد الفرض البديل $\mu \neq 75$ عند مستوى معنوية $\alpha = 0.05$. اختبرت عينة عشوائية من الحجم $n = 15$ وكان متوسط درجات الطلبة في العينة $\bar{x} = 80$. ما الاستنتاج الذي يمكن وضعه ؟

١٣- تسلم أحد التجار كمية كبيرة من بطاريات السيارات المنتجة بواسطة مصنع جديد. يعتقد مدير المصنع أن متوسط عمر البطارية المنتجة 30 شهرا . لاختبار ذلك اختبرت عينة عشوائية من البطاريات تسلمها التاجر وتمت تجربتها فكانت أعمارها بالشهر كالآتي: 29.9, 30.0, 30.2, 35.4, 37.6, 37.6, 92.6, 29.8, 34.7, 39.8, 35.4. فإذا كانت أعمار البطاريات المنتجة في المصنع تتبع توزيعاً طبيعياً فهل تدل بيانات العينة على أن متوسط أعمار البطاريات أقل من 30 شهر إذا علم أن الانحراف المعياري للبطاريات المنتجة في هذا المصنع $\sigma = 3$ شهرا .

١٤- إذا كان متوسط الذكاء لعينة عشوائية ، من الحجم $n = 20$ ، هو $\bar{x} = 1053$. وبفرض أن درجات الذكاء في المجتمع الذي اختبرت فيه العينة تتبع توزيعاً طبيعياً بانحراف معياري 3.7. تحقق من صحة الفرض القائل أن متوسط ذكاء المجتمع الذي اختبرت منه العينة لا يختلف عن 106.2 عند مستوى معنوية $\alpha = 0.05$.

١٥- اختبرت عينة عشوائية من 25 عاملاً بإحدى الشركات وكان متوسط إنتاج العامل في العينة 28 وحدة في اليوم، علماً بأن إنتاجية العامل في هذه الشركة تتبع توزيعاً طبيعياً بـ $\sigma^2 = 5$ اختبر الفرض القائل أن متوسط إنتاجية العامل اليومية في هذه الشركة 25 وحدة وذلك عند مستوى معنوية $\alpha = 0.01$.

١٦- إذا كان من المعروف أن كمية المحصول بالمنتج من إحدى المحاصيل تتبع توزيعاً طبيعياً بانحراف معياري 30 إردب للفدان ، اختبرت عينة عشوائية مساحتها 15 فداناً وتم حساب متوسط المحصول في العينة فكانت 490 إردب للفدان اختبر الفرض القائل أن الكمية المتوقعة للمحصول تساوي 500 إردب للفدان وذلك عند مستوى معنوية $\alpha = 0.05$.

١٧- يدعى مستول في مصنع لإنتاج المصابيح الكهربائية أن متوسط عمر المصباح هو 1000 ساعة، مع العلم أن أعمار المصابيح من إنتاج هذا المصنع تتبع توزيعاً طبيعياً بانحراف معياري 30 ساعة . اختبرت عينة عشوائية من الحجم $n = 25$ من إنتاج هذا المصنع فكان $\bar{x} = 900$ فهل تدل هذه البيانات أن متوسط عمر المصابيح 1000 ساعة وذلك عند مستوى معنوية

$$\alpha = 0.01$$

١٨- تستخدم آلة للقياس بسلمة ما بطريقة أوماتيكية بموسط 55 جرام للكيس. اختبرت عينة عشوائية من 35 كيس من هذه الآلة وتم حساب متوسط وزن الكيس فكان $\bar{x} = 57$ والانحراف المعياري $s = 3$ جرام اختبر فرض العدم $H_0: \mu = 55$ ضد الفرض البديل $H_1: \mu \neq 55$ وذلك عند مستوى معنوية $\alpha = 0.05$.

١٩- للمعلومات التالية إذا علم أن العينة تم اختيارها من مجتمع له توزيعاً طبيعياً، أوجد منطقة الرفض وأحسب قيمة z وقرر هل تقبل فرض العدم أم لا ؟

$H_0: \mu = 30, H_1: \mu < 30, \sigma = 6, n = 20, \alpha = 0.05, \bar{x} = 25$.

٢٠- يعتقد مسئول في دار للنشر أنه في المتوسط تطبع الآلة 45 نسخة في الدقيقة من مجلد ما علماً بأن المجتمع الذي اختبرت منه العينة يتبع توزيعاً طبيعياً بانحراف معياري $\sigma = 2.1$ دقيقة. في محاولة لزيادة عدد النسخ التي تطبع تم إدخال تعديلات على آلة الطبع ثم أخذت عينة عشوائية من الحجم $n = 25$ فكان متوسط عدد النسخ في الدقيقة $\bar{x} = 55$ اختبر فرض العدم $H_0: \mu = 45$ ضد الفرض البديل $H_1: \mu > 45$ عند مستوى معنوية $\alpha = 0.05$.

٢١- إذا كان متوسط القامة في مجتمع ما هو 172 سم. اختبرت عينة عشوائية من هذا المجتمع فكان متوسط الطول فيها 173 سم بانحراف معياري 3.5. فإذا كان حجم العينة 30 شخصاً. اختبر فرض العدم $\mu = 172$ ضد الفرض البديل $\mu \neq 172$ عند مستوى معنوية $\alpha = 0.01$.

٢٢- اختبرت عينة عشوائية من 35 أسرة وحسب متوسط الدخل الشهري للأسرة فوجد أنه \$500 بانحراف معياري 40 \$. اختبر فرض العدم أن متوسط دخل الأسرة في هذا المجتمع يساوي \$450 ضد الفرض البديل $\mu \neq 450$ وذلك عند مستوى معنوية $\alpha = 0.01$.

٢٣- في دراسة لمعرفة تأثير نوع معين من معجون الأسنان على مقاومة تسوس الأسنان اختبرت عينة عشوائية من 25 طالباً وأشرف عليهم أحد المدرسين حتى يستعملوا هذا المعجون يومياً ثم قيست المقاومة بعد ثلاثة أشهر من هذه المحاولة بقياس معين لعدد الميكروبات من نوع معسروف موجود في اللعاب فكانت $\bar{x} = 5$. فإذا كان معروف للباحث أن هذا المقياس يتبع التوزيع الطبيعي بانحراف معياري $\sigma = 0.49$. فهل هناك ما يدعو إلى الاعتقاد بأن هذا المعجون لساندة على زيادة المقاومة ضد التسوس وذلك عند مستوى معنوية $\alpha = 0.01$ ؟ علماً بأن $\mu_0 = 3$.

٢٤- قررت شركة ما لا تزيد مدة المكالمات التلفونية التي يطلبها الموظف عن 25 دقيقة. اختبرت عينة عشوائية من 35 مكالمات فأعطيت متوسط 26.25 دقيقة بانحراف معياري 2.1

دقيقة ما الاستنتاج الذي يمكن اتخاذه بناء على هذه العينة وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٥- اختبرت عينة من 35 عاملا في مصنع وكان متوسط العمر 38 سنة بانحراف معياري 8 سنة. اختبر فرض العدم $\mu = 40$ ضد الفرض البديل $\mu \neq 40$ عند مستوى معنوية $\alpha = 0.05$.

-٢٦- في تقرير خاص وجد أن كمية مركب tar لكل سيجارة من نوع ما هو 13 ملليجرام . اختبرت عينة عشوائية من 100 سيجارة فأعطيت متوسط 12.85 من tar بانحراف معياري 0.75 اختبر فرض العدم $\mu = 13$ ضد الفرض البديل $\mu \neq 13$ عند مستوى معنوية $\alpha = 0.01$.

-٢٧- قام مسئول من مصنع للعب الأطفال بعمل تعديل في سيارات الأطفال وكان من نتيجة هذا التعديل زيادة زمن عمل البطارية المشغلة للسيارة اختبرت عينة عشوائية من 100 سيارة وتم حساب متوسط عمر البطارية فكان 31 ساعة بانحراف معياري 0.75 اختبر فرض العدم $\mu = 25$ وذلك عند مستوى معنوية $\alpha = 0.01$.

-٢٨- إذا كانت نقطة اللهبان الخسوبة من 16 عينة من نوع ما من زيت الطعام المهلج هي $\bar{x} = 94.32$ بفرض أن توزيع نقطة اللهبان طبيعي بانحراف معياري $\sigma = 1.2$. المطلوب اختبار فرض العدم $\mu = 95$ ضد الفرض البديل $\mu \neq 95$ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٩- يعتقد مسئول في مصنع للتلفزيونات أن التيار اللازم للحصول على صورة واضحة لشاشة التلفزيون على الأكثر $250 \mu A$ اختبرت عينة عشوائية من 20 وحدة (تلفزيون) وكان متوسط التيار من العينة $\bar{x} = 257.3$ فإذا كان μ تمثل متوسط التيار الحقيقي الضروري للحصول على الصورة الواضحة للتلفزيون من هذا النوع وتحت فرض أن μ هي متوسط توزيع طبيعي بانحراف معياري $\sigma = 15$ ، اختبر فرض العدم $\mu = 250$ ضد الفرض المناسب وذلك عند مستوى معنوية $\alpha = 0.01$.

-٣٠- إذا كان T متغيرا عشوائيا يتبع توزيع t عندما يكون H_0 صحيحا أوجد مستوى المعنوية α لكل من الحالات الآتية : (درجات الحرية = v) .

$$H_1: \mu > \mu_0 , \quad v = 15 \quad \text{أ-}$$

$$\text{منطقة الرفض } T \geq 3.733$$

$$H_1: \mu < \mu_0 , \quad n = 24 \quad \text{ب-}$$

ومنطقة الرفض $T \leq -2.069$

$$\text{ج- } H_1: \mu \neq \mu_0, n=31$$

ومنطقة الرفض $T > 1.697$ $T \leq -1.697$

٣١- يقوم باحث بجمع البيانات لاختبار فرض العدم $H_0: \mu = 17$ ضد الفرض البديل $H_1: \mu > 17$ أوجد مستوى المعنوية α المرتبط بقيم t ودرجات الحرية التالية :

$$\text{أ- } v=14, t=1.761$$

$$\text{ب- } v=25, t=3.450$$

$$\text{ج- } v=13, t=1.771$$

$$\text{د- } v=8, t=1.860$$

$$\text{هـ- } v=40, t=1.684$$

٣٢- أوجد مستوى المعنوية α لاختبار من جانبيين في الحالات التالية :

$$\text{أ- } v=6, t=2.447$$

$$\text{ب- } v=14, t=2.145$$

$$\text{ج- } v=24, t=2.064$$

$$\text{د- } v=17, t=2.110$$

٣٣- اختبرت عينة عشوائية مكونة من درجات 10 طلاب في مادة الإحصاء وكان $s=3$, $\bar{x}=7$. اختبر الفرض القائل أن متوسط درجة الطالب في المجتمع الطبيعي المسحوب منه العينة يساوي 7.5 وذلك عند مستوى معنوية $\alpha=0.05$.

٣٤- اختبرت عينة عشوائية من 10 فكانت أطوالهم بالبوصة هي

81, 70, 68, 68, 64, 72, 71, 80, 61, 70

اختبر الفرض القائل أن $\mu=75$ ضد الفرض البديل $\mu \neq 75$ عند مستوى معنوية $\alpha=0.01$. مع العلم بأن المجتمع الذي اختبرت منه العينة يتبع توزيعاً طبيعياً.

٣٥- يعتقد مدير شركة لصناعة الصابون أن أوزان صناديق الصابون يتبع التوزيع الطبيعي بمتوسط 500 جرام. اختبرت عينة عشوائية من 15 صندوقاً ووجد أن $\bar{x}=530$ و $s=5.5$ جراماً. اختبر فرض العدم أن $\mu=500$ ضد الفرض البديل $\mu \neq 500$ عند مستوى معنوية $\alpha=0.05$.

٣٦- اختبرت عينة عشوائية من 20 عبوة من مشروب بارد استخدمت آلة لصنعه. فإذا كان متوسط العبوة $\bar{x}=7.5$ أوقية بانحراف معياري 0.47 أوقية. اختبر فرض العدم أن $\mu=7.8$

ضد الفرض البديل $\mu < 7.8$ عند مستوى معنوية $\alpha = 0.01$ تحت فرض أن المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

-٣٧- لاختبار فرض العدم أن متوسط وزن الصندوق من القمح المباع في شركة ما هو 10 كيلو . اختيرت عينة عشوائية من 10 صناديق فكانت النتائج كالتالي :

10.1, 9.5, 10.1, 11.2, 9.9, 8.7, 6.7, 8.1, 9.9, 6.1

استخدم مستوى معنوية $\alpha = 0.01$ لاختبار فرض العدم وذلك تحت فرض أن توزيع الأوزان لصناديق القمح طبيعي .

-٣٨- يرغب باحث في أن تكون حوضه التربة التي سوف يستخدمها في أبحاثه ذات حموضة $\text{PH} = 8.75$. لهذا الفرض قام الباحث بمعالجة التربة بمحلول من 25% من الماء و 75% من effluent وبعد ذلك تم اختبار 5 عينات مشتقة من التربة المعالجة فكان المتوسط الحسابي والانحراف المعياري هما $s = 0.05$, $\bar{x} = 8$ فهل تدل هذه البيانات على أن حموضة التربة تساوي 8.75 عند مستوى معنوية $\alpha = 0.05$ مع العلم أن العينة تم اختيارها من مجتمع طبيعي .

-٣٩- للمثال (٣٨) اختبر فرض العدم $H_0: \sigma = 0.55$ ضد الفرض البديل $H_1: \sigma \neq 0.55$ عند مستوى معنوية $\alpha = 0.05$.

-٤٠- للمثال (٣٧) اختبر فرض العدم $H_0: \sigma = 0.81$ ضد الفرض البديل $H_1: \sigma < 0.81$ عند مستوى معنوية $\alpha = 0.01$.

-٤١- قام المستولون في شركة لإنتاج ملابس الأطفال بإنتاج نوعية من الملابس المقاومة للحريق . للمقارنة بين النوعين تم الحصول على البيانات التالية :

$n_1 = 15, n_2 = 15, \alpha = 0.05, s_1 = 1.5, s_2 = .2$

اختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.01$ مع العلم أن أزمة الحريق للملابس في المصنع تتبع توزيعاً طبيعياً .

-٤٢- أجريت دراسة في إحدى مراكز العلاج الطبيعي على عينة عشوائية من 10 أشخاص ممن يتبعون نظام إنقاص الوزن وقد تم تسجيل مقدار النقص في الوزن لكل شخص في العينة خلال فترة اتباع النظام المتبع لإنقاص الوزن وتم الحصول على البيانات التالية :

14, 5, 5, 11, 12, 17, 7, 3, 4, 9

اختبر فرض العدم $H_0: \sigma^2 = 10$ ضد الفرض البديل $H_1: \sigma^2 > 10$ عند مستوى معنوية $\alpha = 0.01$ وذلك تحت فرض أن المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

٤٣- اختبرت عينة عشوائية من الحجم $n = 10$ من مجتمع طبيعي وكان تباين العينة $s^2 = 24$ اختبر فرض العدم $H_0: \sigma^2 = 23$ حيث الفرض البديل $H_1: \sigma^2 \neq 23$ وذلك عند مستوى معنوية $\alpha = 0.01$.

٤٤- في إحدى مراكز تعليم غير المبصرين يوجد نظامين A, B لتعليم القراءة فإذا كان تباين العينة لمستوى القراءة بالنسبة للدارسين عن طريق النظام A هو $s_1^2 = 1.2$ وذلك من عينة عشوائية حجمها $n = 30$. اختبرت عينة عشوائية أخرى مكونة من 35 فرداً ممن يدرسون باستخدام النظام B فكان تباين العينة $s_2^2 = 1.04$. اختبر الفرض القائل بأن التدريس باستخدام النظام A متساوي في التشتت مع النظام B عند مستوى معنوية $\alpha = 0.01$.

٤٥- في عينة عشوائية حجمها $n = 20$ وجد أن الانحراف المعياري لتركيز الصوديوم في الدم (mEq/L) هو $s_1 = 40.5$ بينما في عينة أخرى من الحجم $n_2 = 20$ وجد أن $s_2 = 32.1$. بفرض أن توزيع كمية الصوديوم طبيعي اختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.01$.

٤٦- يعتبر التوكسالفين من المبيدات التي تلوث البيئة ويعتبر خطراً على النبات والحيوان والإنسان. قام باحث باختيار عينة عشوائية من الفئران (الإناث) من الحجم $n_1 = 30$ وتم إعطائهم خلطة غذائية بها جرعة منخفضة من التوكسالفين (4 ppm) وأخذت عينة أخرى من الحجم $n_2 = 23$ (اعتبرت عينة المراقبة لم تأخذ أى مبيد في الخلطة الغذائية). وفي نهاية التجربة تم تسجيل الزيادة في الوزن لكل فأر، فإذا كان الانحراف المعياري لعينة المراقبة $s_1 = 32$ جراماً بينما الانحراف المعياري للعينة المعالجة $s_2 = 54$ جراماً اختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.01$ ، مع العلم أن العنيتين تم اختيارهما من مجتمعين طبيعيين.

٤٧- اختبرت 10 عينات من الموقع A و 8 عينات من الموقع B وتم قياس الحموضة لكل عينة PH. والبيانات في الجدول التالي:

الموقع A	7.93	7.88	8.01	8.01	8.52	8.53
			7.80	7.92	7.85	9.98
الموقع B	3.35	7.4	7.58	7.73	7.85	
			7.27	7.27		7.3

أ- اختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- هل هناك فرق معنوي بين متوسطي الحموضة للموقعين عند مستوى معنوية $\alpha = 0.02$ مع العلم أن العنيتين تم اختيارهما من مجتمعين طبيعيين .

-٤٨- زرع صنفان من القمح ، الأول من 10 قطع فكان تباين الغصول هو $s_1^2 = 25$ وزرع الصنف الثاني في 15 قطعة وكان تباين الغصول هو $s_2^2 = 15$ اختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$ مع العلم أن العنيتين تم اختيارهما من مجتمعين طبيعيين .

-٤٩- إذا كانت μ_1 تمثل العمر الحقيقي لإطارات السيارات من النوع A مفاصة بالأميال (عدد الأميال التي تقطعها السيارة حتى يستهلك الإطار) و μ_2 تمثل العمر الحقيقي لإطارات السيارات من النوع B . اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ إذا كانت :

$$n_1 = 40, \bar{x}_1 = 36500, s_1 = 220,$$

$$n_2 = 40, \bar{x}_2 = 33400, s_2 = 190$$

مع العلم بأن العنيتين تم اختيارهما من مجتمعين طبيعيين .

-٥٠- طبق اختبار للمصايبه على مجموعتين ، الأولى من الذكور وحجمها 35 والثانية من الإناث وحجمها 40 فإذا كان متوسط المصايبه لدى الذكور 21.3 بانحراف معياري 4.6 ومتوسط المصايبه لدى الإناث 24.2 بانحراف معياري 3.9 تحقق من صحة الفرض القائل أن $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$.

-٥١- في دراسة لتقدير ما إذا كان هناك فرق معنوي بين أجور أعضاء هيئة التدريس في جامعتين A و B ، اختبرت عينة عشوائية من 100 عضو هيئة تدريس من الجامعة A ووجد أن $\bar{x}_1 = 11000$ \$ خلال ٩ أشهر بانحراف معياري $s_1 = 1300$. كما اختبرت عينة عشوائية أخرى من 200 عضو هيئة تدريس من الجامعة B ووجد أن $\bar{x}_2 = 11900$ \$ بانحراف معياري $s_2 = 1400$. اختبر الفرض القائل أن متوسط الأجور لأعضاء هيئة التدريس خلال 9 أشهر في الجامعة A لا يختلف عن متوسط الأجور في الجامعة B عند مستوى معنوية $\alpha = 0.05$.

-٥٢- طبق اختبار لطلاقة الكلمات على مجموعتين الأولى من الانبساطين والثانية من العصابين حجمهما 45 ، 50 شخصا على التوالي ، وحصلت مجموعة الانبساطين على متوسط قدرة 70.2 بانحراف معياري 11.6 وحصلت العصابين على متوسط قدرة 65 بانحراف معياري 10.2 . تحقق من صحة الفرض القائل أن طلاقة الكلمات واحدة في المجموعتين؟

-٥٣- يعتقد المسئول في مصنع عن وجود اختلاف في جودة نوعين من قطع الفيار تم استلامهم من موردين A, B, وقد تم الحصول على البيانات التالية بناء على عيتين عشوائيتين من الموردين :

$$\text{المورد A} \quad n_1 = 50, \quad \bar{x}_1 = 153, \quad s_1 = 10$$

$$\text{المورد B} \quad n_2 = 100, \quad \bar{x}_2 = 150, \quad s_2 = 5$$

اختبر القول القائل بعدم وجود فرق معنوي بين العيتين عند مستوى معنوية $\alpha = 0.05$.

-٥٤- في إحدى مراكز رعاية الطفل سجلت الأوزان (بالرطل) لعيتين عشوائيتين من الأطفال حديثي الولادة كل منهما يتكون من خمسة أطفال فكانت البيانات كإيلي :

$$\text{المجموعة الأولى} \quad 8, 6, 6, 5, 7$$

$$\text{المجموعة الثانية} \quad 8, 7, 8, 5, 6$$

أ- اختبر الفرض القائل أن $\sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $\sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.1$. مع العلم أن العيتين تم اختيارهما من مجتمعين طبيعيين.

ب- اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.02$.

-٥٥- يرغب مسئول في مصنع لإنتاج معجون للأسنان في دراسة تأثير إضافة مادة كيميائية معينة إلى معجون لتحسين مفعوله. اختبرت عيتين مستقلتين كل عينة من 10 أشخاص استخدمت أفراد العينة الأولى المعجون مضاف إليه المادة الكيميائية بينما استخدمت أفراد العينة الثانية المعجون بدون إضافة المادة الكيميائية باستخدام مقياس خاص تم الحصول على البيانات التالية :

$$\text{العينة الأولى (أضيفت المادة الكيميائية)} \quad \bar{x}_1 = 8, \quad s_1 = 3$$

$$\text{العينة الثانية (بدون إضافة)} \quad \bar{x}_2 = 9, \quad s_2 = 4$$

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.01$. مع العلم أن العيتين تم اختيارهما من مجتمعين طبيعيين.

-٥٦- مجموعتان من الطلبة تتكون إحداهما من 10 أفراد والأخرى من 8 أفراد أعطيت امتحانا واحدا ودونت النتائج كما يلي :

$$\text{المجموعة الأولى} \quad 7, 7, 7, 6, 7, 8, 8, 9, 9, 10$$

$$\text{المجموعة الثانية} \quad 10, 10, 8, 7, 7, 10, 9, 7$$

المطلوب اختبار الفرض $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.1$. مع العلم أن العنيتين تم اختيارهما من مجتمعين طبيعيين .

٥٧- في اختبار للفرق لمجموعة من الذكور والإناث تم الحصول على البيانات التالي :

	الانحراف المعياري	متوسط العينة	حجم العينة
ذكور	4.83	10.4	20
إناث	4.68	9.26	25

اختبر فرض العلم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ مع العلم بأن العنيتين تم اختيارهما من مجتمعين طبيعيين .

٥٨- البيانات التالية تمثل أوزان الأدوات الخاصة بأعضاء ناديين A , B (الوزن بالكيلو) :

النادي A	32	28	41	39	34
النادي B	36	39	30	34	35

اختبر فرض العلم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ مع العلم بأن العنيتين تم اختيارهما من مجتمعين طبيعيين .

٥٩- تعتقد اللباسة الحديثة أن عريجي المعاهد المتوسطة يتزوجون في عمر أقل من عريجي الجامعات. لتدعيم هذا الاعتقاد ، اختبرت عينة عشوائية من 100 فرد من كل مجموعة وتم تسجيل العمر عند الزواج وكانت النتائج كما يلي :

عريجي المعاهد المتوسطة	$\bar{x}_1 = 22.5$, $s_1 = 1.4$
عريجي الجامعات	$\bar{x}_2 = 28$, $s_2 = 1.5$

اختبر فرض العلم بعدم وجود فرق معنوي بين المجموعتين وذلك عند مستوى معنوية $\alpha = 0.01$.

٦٠- تم تقسيم مجموعة من الأطفال حديثي الولادة في مستشفى إلى مجموعتين ، كل مجموعة استخدمت نوع من لبن الأطفال وقد تم تسجيل وزن الأطفال في كل مجموعة وذلك بعد 6 أسابيع من الولادة ، كانت النتائج كما يلي بالكيلو :

المجموعة A	5.6	6.1	4.6	5.2	5.0	4.5	4.2	3.0
المجموعة B	8.6	8.7	5.8	5.5	4.4	4.5	4.2	

المطلوب معرفة هل هناك فرق معنوي بين النوعين عند مستوى معنوية $\alpha = 0.1$

علما بأن العنيتين تم اختيارهما من مجتمعين طبيعيين .

٦١- المقارنة إحدى العناصر المعدنية لنوعين من العصار A , B أخذت عينة عشوائية من العلب المعروضة في الأسواق لكل منهما وكانت النتائج كما يلي (وحدة القياس mg/100

(gm

النوع A	4.6	4.2	4.5	4.3	4.9	5.2	5.7
النوع B	5.1	4.3	4.2	5.7	5.8	5.1	5.3

أ- أختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- أختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.02$ وذلك تحت فرض أن العنيتين تم اختيارهما من مجتمعين طبيعيين .

٦٢- أعطى اختبار في القراءة لعينة من 12 طفل من أصل أمريكي وعينة أخرى من 10 أطفال من أصل مكسيكي . نتائج هذا الاختبار كانت :

$$\bar{x}_1 = 74 , s_1 = 8 \quad \text{أمريكي الأصل}$$

$$\bar{x}_2 = 70 , s_2 = 10 \quad \text{مكسيكي}$$

أ- أختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- أختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$ وذلك تحت فرض أن العنيتين تم اختيارهما من مجتمعين طبيعيين .

٦٣- لدراسة تأثير نوع من السماد على مساحة الورقة لنبات ما . اختبر 16 قطعة للزراعة وتم اختبار 8 منهم عشوائيا للمعالجة بالسماد والباقي تركت لمعالجة المراقبة وتم زراعة النباتات في القطع التي عددها 16 . البيانات التالية تعطى مساحة الورقة لكل نبات .

المعالجة بالسماد	1024	1216	1312	1780	
	1216	1312	992	1120	
معالجة المراقبة	1104	1072	1088	1328	1376
	1280	1120	1200		

أختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ علما بأن العنيتين تم اختيارهما من مجتمعين طبيعيين .

٦٤- تعطى النتائج التالية الزيادة في الوزن (gm) لعنيتين الأولى عينة مراقبة والأخرى تم معالجة أفرادها بمركب Soft steroid (الجرعة 1 mg / pellet dose)

	n_i	\bar{x}_i	s_i
المراقبة	10	40.5	2.5
Soft Steroid	8	32.8	2.6

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.1$ علما بأن العنيتين تم اختيارهما من مجتمعين طبيعيين .

-٦٥- في دراسة لأهمية دخول الطالب المعمل في مادة الفيزياء اختبرت عينة عشوائية من 11 طالبا تم دخولهم المعمل. وكان متوسط الدرجة النهائية $\bar{x}_1 = 85$ بانحراف معياري $s_1 = 4.6$ كما اختبرت عينة أخرى من 17 طالبا من الذين لم يدخلوا المعمل فكان $\bar{x}_2 = 79$ بانحراف معياري $s_2 = 6.2$. اختبر الفرض القائل أن دخول المعمل يؤثر على درجات الطلبة عند مستوى معنوية $\alpha = 0.1$ وذلك تحت فرض أن العنيتين تم اختيارهما من توزيعين طبيعيين .

-٦٦- إذا كانت نسبة الألياف في عتين عشوائيتين من شجيرات الكتان هي :

العينة الأولى	11.9	11.8	12.4	12.6	11.9	11.2
العينة الثانية	12.9	12.6	12.7	12.8	12.7	

المطلوب اختبار ما إذا كانت العنيتين مأخوذتان من مجتمعين هما نفس المتوسط والانحراف المعياري عند مستوى معنوية $\alpha = 0.1$ وذلك تحت فرض أن العنيتين تم اختيارهما من مجتمعين طبيعيين .

-٦٧- البيانات التالية تمثل أزمنة العرض بالدقيقة لأفلام منتجة من شركتي A, B.

الشركة A	102	88	98	109	92	
الشركة B	81	166	98	93	88	110

اختبر الفرض القائل أن متوسط زمن العرض للفيلم المنتج من الشركة A يزيد عن متوسط زمن العرض للفيلم المنتج من الشركة B وذلك تحت فرض أن توزيع الزمن للشركتين طبيعي (مستوى المعنوية $\alpha = 0.1$) .

-٦٨- يعتقد مدير شركة ما أن الأجر في الساعة بالدولار للعمال النصف ماهرة متساوية في قسمين من الشركة كالآتي :

القسم الأول :	$\bar{x}_1 = 5.3,$	$n_1 = 50$	$s_1 = 2.1$
القسم الثاني :	$\bar{x}_2 = 5.9,$	$n_2 = 40$	$s_2 = 1.7$

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ عند مستوى معنوية $\alpha = 0.1$.

٦٩- في مقارنة لم توسط درجات الذكاء IQ لفرقتين في كلية ما اختبرت عينة عشوائية من أربعة طلبة من إحدى الفرق وكانت درجات الذكاء لهم 110, 113, 116, 117. كما اختبرت عينة عشوائية من الفرقة الأخرى وكانت درجات الذكاء لهم 109, 111, 112, 112. اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 > \mu_2$ عند مستوى معنوية $\alpha = 0.1$ تحت فرض أن العنيتين تم اختيارها من توزيعين طبيعيين .

٧٠- اختبرت عينة عشوائية من البحارة حجمها $n = 10$ وقيست أوزانهم بالبوصة فكانت

80, 70, 68, 66, 72, 73, 70, 61, 60, 72

كما اختبرت عينة عشوائية أخرى من الجنود فكانت أوزانهم كيلي :

61, 61, 72, 70, 73, 62, 63, 71, 70

أ- اختبر فرض العدم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.01$ تحت فرض أن العنيتين تم اختيارها من توزيعين طبيعيين .

٧١- تم قياس كمية الأنسولين لعينة من الفئران المصابة بمرض السكر والتي تصالغ بجرعات منخفضة من الأنسولين (المجموعة A) وأخرى تعالج بجرعات عالية من الأنسولين (المجموعة B) وتم الحصول على البيانات التالية :

المجموعة A $n_1 = 8, \bar{x}_1 = 1.98, s_1 = 0.5,$

المجموعة B $n_2 = 12, \bar{x}_2 = 1.3, s_2 = 0.35$

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.01$ تحت فرض أن العنيتين تم اختيارها من توزيعين طبيعيين .

٧٢- البيانات التالية تمثل عدد البكتريا الهوائية (عدد المستعمرات / قدم 3) والمأخوذة من 8 عينات مأخوذة من حجرات مفروشة بالسجاد و 8 عينات غير مفروشة بالسجاد في مستشفى ما

14 14.6 10.1 10.8 13 7.1 8.2 11.2 حجرات مفروشة

13.7 10.1 11.1 12 7.2 3.8 8.2 12.1 حجرات غير مفروشة

اختبر فرض العدم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.02$ تحت فرض أن العنيتين تم اختيارها من توزيعين طبيعيين .

٧٣- إذا كانت عدد البكتريا بالملايين على نبات محفوظ على درجتي 10° , 20° بعد 10 أيام هو :

10 9 8 7 7 6 9 8 10°

20 6 5 3 2 1 1 3 4

أ- اختبار فرض العلم $H_0: \sigma_1^2 = \sigma_2^2$ ضد الفرض البديل $H_1: \sigma_1^2 \neq \sigma_2^2$ عند مستوى معنوية $\alpha = 0.02$.

ب- اختبار فرض العلم $H_0: \mu_1 = \mu_2$ ضد الفرض البديل $H_1: \mu_1 \neq \mu_2$ عند مستوى معنوية $\alpha = 0.05$ تحت فرض أن العتتين تم اختيارهما من توزيعين طبيعيين.

٧٤- طبق اختبار للقدرة على التفكير الناقد على مجموعة من المراهقين قبل حضورهم برنامج أعد لهذا الغرض مدته 40 أسبوعا وبعد حضورهم للبرنامج فإذا كان حجم العينة 10 أفراد تحقق من صحة الفرض القائل أن للبرنامج فاعلية على تنمية الفكر الناقد لدى المراهقين إذا كانت البيانات كالتالي :

الشخص	1	2	3	4	5	6	7	8	9	10
قبل البرنامج	15	15	11	10	10	15	17	16	11	12
بعد البرنامج	20	25	14	16	22	23	24	25	24	26

وذلك عند مستوى معنوية $\alpha = 0.05$.

٧٥- مقارنة علقيتين من ناحية تأثيرهما على نمو العجول خلال شهرين من التغذية المختصت عينة عشوائية من 10 أزواج (توائم) من العجول. اختبار صحة الفرض القائل $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.01$. إذا توافرت لك البيانات التالية :

التوائم	1	2	3	4	5	6	7	8	9	10
العليقة أ	20.1	30.6	31.7	27.2	22.7	22.4	29.3	28.8	26.3	27.1
العليقة ب	25.5	32.2	30.1	29.3	25.1	30.1	30.6	28.1	30.2	30.1

٧٦- لدراسة الآثار الجانبية لدواء معين على ضغط الدم أخذت قياسات ضغط الدم لعينة عشوائية من المرضى حجمها 8 قبل وبعد تناول الدواء فكانت البيانات كالتالي :

الشخص	1	2	3	4	5	6	7	8
قبل الدواء	70	72	80	65	66	77	70	72
بعد الدواء	66	70	75	62	64	73	74	77

اختبر الفرض القائل $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

٧٧- طبق مقياس للاتجاهات نحو الأطفال في عينة من السيدات الآتي تم قبولهم بقسم دراسات الطفولة فور التحاقهن بالقسم ، ثم طبق نفس المقياس مرة أخرى فور حصولهن على البكالوريوس وكانت النتائج كالتالي :

السيدة	1	2	3	4	5	6	7	8	9	10
عدد الالتحاق	9	6	6	14	14	8	8	12	8	14
عدد التخرج	11	20	15	15	15	18	18	17	16	16

اختبر الفرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

٧٨- قام شخص بإجراء 6 عمليات حياية على آلتين حاسبتين وتسجيل الزمن اللازم لكل عملية على كل آلة في الجدول التالي :

العملية الحسابية	1	2	3	4	5	6
الآلة الأولى	22	17	28	21	32	19
الآلة الثانية	18	19	23	22	30	22

اختبر الفرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

٧٩- للمقارنة بين طريقتين لتقدير كمية اللبن المهضوم في الأطفال الرضع ، اختبرت عينة عشوائية من الحجم $n = 14$ البيانات التالية تعطى كمية اللبن المهضوم لكل طفل في العينة :

الطفل	1	2	3	4	5	6	7
Isotopic	1509	1418	1561	1556	2169	1760	1098
Test-weighing	1448	1254	1336	1565	2000	1318	1410
الطفل	8	9	10	11	12	13	14
Isotopic	1198	1479	1781	1414	1954	2174	2058
Test-weighing	1179	1342	1124	1468	1604	1722	1518

اختبر الفرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

٨٠- للمقارنة من نوعين من مصائد الأسماك الخاصة بسمك الماهور ، استخدم النوعين خلال 16 فترة زمنية خلال 4 سنوات وذلك في بحيرة ما و البيانات التالية تعطى كمية الأسماك لكل يوم

الفترة	1	2	3	4	5	6	7	8
Pipe	6.64	7.89	1.83	0.42	0.85	0.29	0.57	0.63
Brush	9.73	8.1	2.17	0.75	1.61	0.75	0.83	0.56
الفترة	9	10	11	12	13	14	15	16
Pipe	0.32	0.37	0.0	0.11	4.86	1.8	0.32	0.88
Bursh	0.67	0.32	0.98	0.52	5.38	2.33	0.91	0.70

اختبر الفرض العلم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$

٨١- من المعروف أن العناصر النادرة لها تأثير غير مرغوب فيه على طعم مياه الشرب وأيضا آثار ضارة على الصحة إذا وجدت بكميات كبيرة . قام باحث بإعبار الزنك (مقاس mg/L) لكل من سطح وقاع النهر في كل موقع . الأزواج الستة من المشاهدات معطاة في الجدول التالي . هل تدل هذه البيانات على أن المتوسط الحقيقي لتركيز الزنك في قاع الماء يزيد عن سطح الماء ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الموقع	1	2	3	4	5	6
قاع النهر	0.430	0.266	0.567	0.531	0.707	0.716
سطح النهر	0.415	0.238	0.390	0.410	0.605	0.609

٨٢- لاختبار تأثير إضافة مواد الجفاف على الدهان أجريت تجربة على 6 عينات من المعدن حيث قسمت كل عينة إلى نصفين لكل عينة ثم دهان نصف العينة بالدهان المضاف له مواد الجفاف بينما تم دهان النصف الآخر من العينة بالدهان بدون إضافة مواد الجفاف . وترك العينات التي علدها 12 حتى الجفاف وتم تسجيل زمن الجفاف في الجدول التالي :

العينة	1	2	3	4	5	6
الدهان مضاف له مواد الجفاف	3.4	3.8	4.2	4.1	3.5	4.7
الدهان بدون إضافة	3.6	3.8	4.3	4.3	3.6	4.6

اختبر الفرض العلم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ وذلك عند مستوى

معنوية $\alpha = 0.05$

٨٣- البيانات التالية تمثل أوزان 10 قوائم عند لحظة الميلاد (مقاس بالكيلو) وذلك عند مستوى معنوية $\alpha = 0.05$

القوائم	1	2	3	4	5	6	7	8	9	10
الأول في الولادة	3.95	3.41	3.73	3.48	4.28	3.98	4.18	4.04	3.73	4.13
الثاني في الولادة	3.93	3.35	3.72	4.18	3.44	4.15	3.89	4.20	4.00	3.72

اختر فرض العدم أن الطفل الأول في الولد لأي قوائم أقل في الوزن من الطفل الثاني .

٨٤- إذا كان من المعلوم أن واحد من كل سبعة يدخنون في إحدى الدول . فإذا أجريت حملة للتوعية عن مضار التدخين . للحكم على مدى نجاح تلك الحملة أخذت عينة عشوائية من 600 شخص ووجد من بينهم 180 لا يزالون يدخنون هل البيانات تعطي دليلاً كافياً على نجاح الحملة وذلك عند مستوى معنوية $\alpha = 0.05$.

٨٥- أصيبت إحدى المناطق الزراعية التي مساحتها 6000 فدان بإحدى الآفات الزراعية . اختبرت عينة عشوائية حجمها 300 فدان فوجد أن نسبة الإصابة بها 20% اختبر الفرض القائل أن نسبة الإصابة 15% عند مستوى معنوية $\alpha = 0.01$.

٨٦- أجريت دراسة على المركز المالي لعينة عشوائية حجمها 200 من شركات الغزل والنسيج فوجد أن 25% منها سوف تتجه إلى الإفلاس بعد عام . اختبر الفرض القائل أن 40% من شركات الغزل والنسيج في المجتمع المسحوب منه العينة سوف تتجه إلى الإفلاس بعد عام وذلك عند مستوى معنوية $\alpha = 0.01$.

٨٧- اختبرت عينة عشوائية من 200 شخص من مجتمع ما ووجد أن 40 شخص من العينة مصابون بمرض ما . المطلوب اختبار الفرض أن نسبة الإصابة بالمرض في هذا المجتمع أقل من 0.5 وذلك عند مستوى معنوية $\alpha = 0.01$.

٨٨- يعتقد مدير الإنتاج في مصنع لإنتاج التلفزيونات في بلد ما أن 80% من الأسر تمتلك تلفزيون ملون . للتحقق من هذا الفرض اختبرت عينة عشوائية من 1000 أسرة ووجد أن 318 منهم يمتلكون تلفزيوناً ملوناً اختبر صحة هذا الفرض ($p = 0.8$) عند مستوى معنوية $\alpha = 0.05$.

- ٨٩- يعتقد المستولن في كلية ما أن 25% من الطلبة يمتلكون سيارة . اختبر صحة هذا الفرض إذا اختبرت عينة عشوائية من 90 طالبا ووجد أن 28 منهم يمتلك سيارة وذلك عند مستوى معنوية $\alpha = 0.05$.
- ٩٠- في دراسة لتقدير نسبة ربات البيوت اللاتي يمتلكن غسالة بمجفف في بلدنا ما ، اختبرت عينة عشوائية من 1000 ربة مرل ووجد أن 630 منهم يمتلكن غسالة بمجفف . اختبر صحة الفرض القائل أن $p = 0.4$ ضد الفرض البديل $p \neq 0.4$ عند مستوى معنوية $\alpha = 0.05$.
- ٩١- في تقرير خاص وجد أن 30% فقط من خريجي الجامعات يجدون عملا في مجال تخصصهم الرئيسي اختبرت عينة عشوائية من 500 خريج فوجد أن بينهم 250 يعملون في مجال تخصصهم اختبر فرض العدم $H_0: p = 0.3$ ضد الفرض البديل $H_1: p \neq 0.3$ عند مستوى معنوية $\alpha = 0.01$.
- ٩٢- يعتقد أن نسبة الأسر التي تشتري اللبن من المصنع A في مدينة ما هو $p = 0.6$ فإذا اختبرت عينة عشوائية من 200 أسرة ووجد أن نسبة الأسر التي تشتري اللبن 80. اختبر لفرض العدم $p = 0.6$ عند مستوى معنوية $\alpha = 0.05$.
- ٩٣- في عينة عشوائية من 400 ناخب في مدينة ما وجد أن 200 منهم يوافقون على فرض 3% ضريبة مبيعات . اختبر فرض العدم أن $p = 0.45$ ضد الفرض البديل $p \neq 0.45$ عند مستوى معنوية $\alpha = 0.01$.
- ٩٤- يعتقد أن 40% من الطلبة في كلية ما يستخدمون نظارة طبية اختبرت عينة عشوائية من 64 طالبا ووجد أن 40 منهم يستخدم نظارة طبية اختبر فرض العدم أن $p = 0.4$ ضد الفرض البديل $p \neq 0.4$ عند مستوى معنوية $\alpha = 0.01$.
- ٩٥- تريد شركة للتليفونات أن تقرر فيما إذا كان بعض الخطوط الجديدة يمكن تجهيزها تحت الأرض. ولأن الرسوم المضافة إلى فاتورة التليفون قليلة بالنسبة إلى التكاليف الباهظة التي تتحملها الشركة لذلك قررت الشركة عمل استفتاء للعملاء وإذا كان 60% من العملاء يفضلون تركيب تحت الأرض فإنها سوف توافق على تجهيز الخطوط الجديدة تحت الأرض وغير ذلك ترفض فإذا كان 118 من العملاء وافق على التركيب من بين 160 عميل اختبر فرض العدم $H_0: p = 0.6$ ضد الفرض البديل $H_1: p \neq 0.6$ عند مستوى معنوية $\alpha = 0.05$.
- ٩٦- في دراسة عن نسبة الكوابيس الليلية اختبرت عينة عشوائية من 160 رجلا وعينة عشوائية أخرى من 192 سيدة وتم سؤلهم على عدد الكوابيس التي تعرضوا لها (على الأقل

خلال شهر) فوجد أن 55 من الرجال و 60 سيدة تعرضوا للكوابيس اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-٩٧- في عينة عشوائية من 300 شاب من المدينة A وجد أن 63 منهم يفضلون قيادة سيارتهم في الطريق الصحراوي على سرعة من 55 , 65 mph بينما وجد أن 75 من 180 في المدينة B يفضلون السرعة من 55 إلى 65 mph . اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-٩٨- في عينة عشوائية من 5726 رقم تليفون في منطقة ما في مارس سنة 1992 ووجد أن 1105 غير مقيدون في الدليل وبعد سنة ومن عينة عشوائية من 5384 ووجد أن 1001 غير مقيدون في الدليل . اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-٩٩- في دراسة عن نسبة الذين يفضلون استخدام حزام الأمان أخذت عينة عشوائية من 200 سائق يعمل على سيارة الغير ووجد أن 115 منهم يفضل استخدام حزام الأمان بينما في عينة أخرى من 300 شخص (يقود سيارته الخاصة) وجد أن 154 يفضل استخدام حزام الأمان . اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.05$.

-١٠٠- ترغب شركة في اختبار ما إذا كانت نسبة المقبول من المكونات الإلكترونية لمورد أجنبي p_1 تزيد عنها لمورد محلي p_2 . اختبرت عينة عشوائية من شحنة كل مورد وتم الحصول على البيانات التالية :

$$n_1 = 200 , n_2 = 200 , \hat{p}_1 = 0.77 , \hat{p}_2 = 0.8$$

اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.01$.

-١٠١- قام أحد العاملين في مكتبة بعمل إحصائية عن مدى تفضيل المترددين للكتب الثقيلة . اختبرت عينة عشوائية من السجلات المترددين على المكتبة فوجد أن 400 من 600 يفضلون قراءة الكتب الثقيلة . كما اختبرت عينة عشوائية من الرجال المترددين على المكتبة فوجد أن 500 من 700 يفضلون قراءة الكتب الثقيلة . اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.01$.

-١٠٢- في دراسة ميدانية لمجموعتان من الأشخاص عن رأيهم في برنامج تليفزيوني المجموعة الأولى مكونة من 600 شخص من الريف بينهم 300 يفضلون البرنامج المجموعة الثانية والمكونة

من 1000 شخص من المدن بينهم 600 شخص يفضلون البرنامج . هل البرنامج شائع في الريف والمدينة بنفس النسبة وذلك عند مستوى معنوية $\alpha = 0.05$.

- ١٠٣ - من مجموعتين متشابهتين ومصابتين بمرض ما ، اختبرت من المجموعة الأولى 50 شخصا ومن المجموعة الثانية 80 شخصا وقد استخدم مع كل عينة نوع مختلف من الدواء فإذا كان عدد الذين تم شفائهم من العينة الأولى 30 وعدد من شفى من المجموعة الثانية 40 فهل هناك دلائل تشير إلى اختلاف نسبة الشفاء باستخدام كل دواء عند مستوى معنوية $\alpha = 0.05$.

- ١٠٤ - في دراسة عن نسبة ربوات البيوت اللاتي يمتلكن غسالة بمجفف في منطقة بين A , B اختبرت عينة عشوائية من الحجم 1000 من المنطقة A فكان عدد اللاتي يمتلكن غسالة 630 . كما اختبرت عينة عشوائية أخرى من المنطقة B حجمها 1200 ووجد أن 710 منهن يمتلكن غسالة بمجفف. اختبر فرض العدم $H_0: p_1 = p_2$ ضد الفرض البديل $H_1: p \neq p_2$ عند مستوى معنوية $\alpha = 0.01$.

الفصل العاشر

الانحدار والارتباط

Regression and Correlation

Regression (١-١٠) الانحدار

يستخدم الانحدار لدراسة العلاقات بين متغيرات قابلة للقياس. تحليل الانحدار له تطبيقات كثيرة في معظم مجالات الدراسة التي تشتمل على العلوم الحيوية ، الفيزيائية ، العلوم الاجتماعية ، الاقتصادية ، الصناعية ... الخ. تختلف الطرق المستخدمة في تحليل الانحدار باختلاف التطبيقات ولكنها تشترك جميعها في الوصف والتنبأ بالقيم المستقبلية . عموما يتكون الانحدار من عدة خطوات. فلدراسة العلاقة بين عدد من المتغيرات ، تجمع البيانات على عينة من المفردات المعرّضة لتلك المتغيرات. في نموذج الانحدار يكون هناك متغير واحد يسمى المتغير التابع **depend variable** أو متغير الاستجابة **response variable** ، بينما المتغيرات الأخرى تسمى متغيرات مستقلة **independent variables** . تستخدم البيانات للحصول على تقديرات لمعالم النموذج . في الحقيقة طرق تقدير معالم النموذج المجهولة كثيرة . سوف تقتصر دراستنا على طريقة المربعات الصغرى **least squares** . يشتمل البند التالي على الانحدار الخطي البسيط ، أي وجود متغير تابع ومتغير مستقل واحد والعلاقة بينهما خطية ، حيث ينحصر اهتمامنا على تعريف النموذج المناسب ومناقشة الفروض وإيجاد تقديرات المربعات الصغرى وعرض بعض طرق التقدير بفترة واختبارات الفروض .

Simple Linear Regression (٢-١٠) الانحدار الخطي البسيط

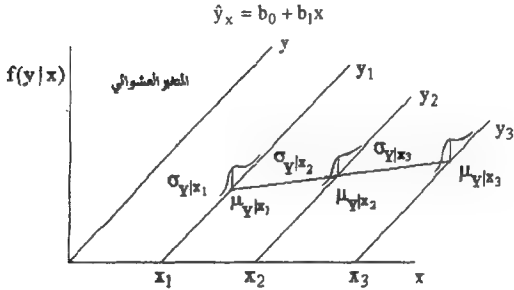
نفرض أننا اخترنا عينة عشوائية من الحجم n من المجتمع موضع الدراسة ممثلة بالنقطة $\{(x_i, y_i) ; i = 1, 2, \dots, n\}$. بتكرار المعاينة مع استخدام نفس قيم x ، فإننا نوقع أن قيم y تختلف من عينة إلى أخرى. وعلى ذلك فإن قيمة y_i في الزوج المرتبة (x_i, y_i) هي قيمة لمتغير عشوائي Y_i . سوف نرمز للتوزيع الاحتمالي للمتغير $Y | x$ المقابل لقيمة ثابتة x بالرمز $f(y | x)$. من الواضح ، أنه إذا كانت $x = x_i$ ، فإن الصيغة $Y | x_i$ تمثل متغير عشوائي Y_i . اهتمامنا سوف يكون في توزيعات فئة المتغيرات العشوائية $\{Y_i, i = 1, 2, \dots, n\}$. والتي يفترض أنها مستقلة. أيضا للحصول على فترات ثقة واختبارات فروض ، لابد أن تكون Y_1, Y_2, \dots, Y_n تتبع توزيعا طبيعيا .

في مشكلة الانحدار سوف نعرف متوسط توزيع Y عند قيمة معطاة x بالرمز $\mu_{Y|x} = E(Y | x)$ وتباين توزيع Y بالرمز $\sigma_{Y|x}^2$ وذلك تحت فرض أن بيانات المتغيرات Y_1, Y_2, \dots, Y_n متساوية ، أي أن $\sigma_{Y|x}^2 = \sigma^2$ ، لجميع قيم x والمعلمة $\mu_{Y|x}$ ثابتة لأي x ولكن قد تختلف باختلاف قيم x .

المنحنى الذي يربط متوسط التوزيعات للمتغيرات Y_1, Y_2, \dots, Y_n يسمى خط الانحدار regression curve . إذا وقعت جميع المتوسطات $\mu_{Y|x}$ على خط مستقيم كما هو موضح في شكل (١٠-١) فإن الانحدار يكون خطي ويمكن تمثيله بالمعادلة :

$$\mu_{Y|x} = \beta_0 + \beta_1 x,$$

حيث β_0 , β_1 معالم النموذج المجهولة والتي يراد تقديرها من بيانات العينة. سوف نرمز لتقديرات هذه المعالم بالرموز b_0 , b_1 على التوالي . وعلى ذلك يمكن تقدير $\mu_{Y|x}$ بواسطة \hat{y}_x من خط الانحدار المقدر التالي :



شكل (١٠-١)

(١٠-٢-١) شكل الانتشار Scatter Plot

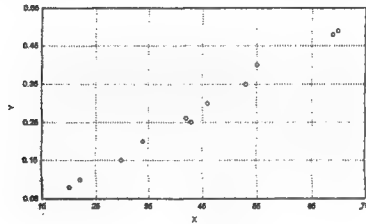
الأسلوب المفيد لبدء تحليل الانحدار البسيط هو تمثيل البيانات بيانياً وهو ما يعرف بشكل الانتشار scatter plot وذلك محاولة اكتشاف الصورة التقريبية للعلاقة . للحصول على شكل الانتشار يخصص محور x (المحور الأفقي) للمتغير المستقل بينما يخصص محور y (المحور الرأسي) للمتغير التابع . لكل زوج (x, y) من أزواج المشاهدات التي عددها n نقوم بتوقيع نقطة على الرسم. كثير من البرامج الإحصائية مثل برنامج Statistica SPSS يمكن استخدامها للحصول على أشكال الانتشار .

مثال (١٠-١) في إحدى التجارب وزن قرون عدد من الغزلان المختلفة الأعمار وكانت النتائج كما هي مدونة في الجدول (١٠-١) . المطلوب رسم شكل الانتشار وتحديد شكل العلاقة بين المتغيرين .

جدول (١٠-١)

المعسر x	20	22	30	34	42	43	46	53	55	69	70
الوزن y	0.08	0.10	0.15	0.20	0.26	0.25	0.30	0.35	0.40	0.48	0.49

الحل . يتضح من شكل (١٠-٢) أن النقط عموماً ، ليس بالضبط ، تقع على خط مستقيم . هذا يجعلنا نقترح أن العلاقة بين المتغيرين يمكن وصفها (كتقريب أولي) بمعادلة خط مستقيم .



شكل (١٠-٢)

Building a Simple Regression Line بناء نموذج الانحدار البسيط (١٠-٢-٢)

بفرض أن كل المتوسطات ، $\mu_{y|x}$ تقع على خط مستقيم وعلى ذلك يمكن كتابة خط المحدار المجتمع على الشكل :

$$\mu_{y|x} = \beta_0 + \beta_1 x$$

كما يمكن كتابة المتغير العشوائي $Y_i = Y | x_i$ على الشكل :

$$Y_i = \beta_0 + \beta_1 x_i + E_i,$$

حيث E_i متغير عشوائي ، بناء على القروض السابقة على Y_i ، من الضروري أن يكون له متوسط صفر وتباين $\sigma^2_{y|x} = \sigma^2$. كل مشاهدة (x_i, y_i) في العينة لا بد أن تحقق العلاقة :

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i.$$

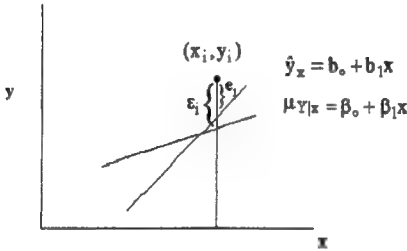
لعينة عشوائية من الحجم n من المشاهدات $\{(x_i, y_i), i=1, 2, \dots, n\}$ فإن خط الانحدار ،
الذي يعتبر تقدير لـ $\mu_{Y|X}$ ، يمكن إيجاد المعادلة :

$$\hat{y}_x = b_0 + b_1 x,$$

وكل زوج من المشاهدات لا بد أن يحقق العلاقة :

$$y_i = b_0 + b_1 x_i + e_i$$

حيث e_i يسمى الباقي residual. الفرق بين e_i و ϵ_i موضح في شكل (٣-١٠).



شكل (٣-١٠)

يعرف مجموع مربعات البواقي كالآتي :

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n [y_i - b_0 - b_1 x_i]^2$$

ويرمز له بالرمز SSE والذي غالباً ما يسمى مجموع مربعات الأخطاء

sum of squares of the errors حول خط الانحدار .

(٣-٢-١٠) طريقة المربعات الصغرى Method of Least Squares

سوف نوجد التقديرين b_0, b_1 للمعالم β_0, β_1 على التوالي بحيث أن مجموع

مربعات الأخطاء يكون أقل ما يمكن. هذه الطريقة لتقدير المعالم تسمى طريقة المربعات الصغرى.

باستخدام حساب التفاضل نجد أن قيمتي β_0 , β_1 اللتين تحققان النهاية الصغرى لمجموع مربعات الأخطاء يمكن تقديرهما بقيمتي b_0 , b_1 اللتين تحققان المعادلتين :

$$b_0 n + b_1 \sum x_i = \sum y_i,$$

$$b_0 \sum x_i + b_1 \sum x_i^2 = \sum x_i y_i$$

وتسميان بالمعادلتين الطبيعيين وبحل هاتين المعادلتين آتيا نحصل على :

$$b_1 = SXY / SXX,$$

$$SXX = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n},$$

$$SXY = \sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n},$$

$$b_0 = \bar{y} - b_1 \bar{x}.$$

مثال (٢-١٠) أجريت تجربة لدراسة العلاقة بين التسميد ومحصول القمح . البيانات التي تم الحصول عليها معطاة في جدول (٢-١٠)

جدول (٢-١٠)

السماد x	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4
محصول y	10	15	30	35	25	30	50	45

(أ) أرسم شكل الانتشار

(ب) أوجد معادلة خط الانحدار المقدرة

الحل . (أ) يتضح من شكل الانتشار (٢-١٠) أن الخط المستقيم هو أفضل طريقة لتمثيل هذه البيانات :

أي أننا نفترض النموذج الخطي البسيط :

(ب) بما أن β_0 , β_1 مجهولتان لأننا نقدرهما من مشاهدات العينة حيث :

$$n = 8 \quad \sum x_i = 10.8 \quad \sum x_i^2 = 18.36$$

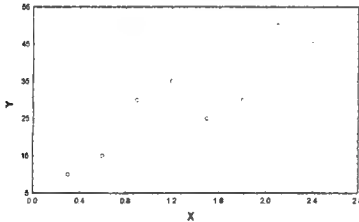
$$\sum x_i y_i = 385.5 \quad , \quad \bar{x} = 1.35 \quad , \quad \bar{y} = 30, \quad \sum y_i = 240.$$

$$\begin{aligned}
 b_1 &= \frac{SXY}{SXX} = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} \\
 &= \frac{385.5 - \frac{(10.8)(240)}{8}}{18.36 - \frac{(10.8)^2}{8}} \\
 &= \frac{61.5}{3.78} = 16.27,
 \end{aligned}$$

$$b_0 = \bar{y} - b_1 \bar{x} = 30 - (16.27)(1.35) = 8.036.$$

معادلة الانحدار المقدرة سوف تكون على الشكل :

$$\hat{y}_x = 8.036 + 16.27 x.$$



شكل (١٠-٤)

Analysis of Variance تحليل الانحدار (١٠-٢-٤)

لاختبار معنوية معامل الانحدار β_1 أي اختبار فرض العدم :

$$H_0 : \beta_1 = 0$$

ضد الفرض البديل :

$$H_1 : \beta_1 \neq 0$$

يجب دراسة مكونات مجموع المربعات وتجزئته إلى مكوناته الأساسية على النحو التالي :

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i)$$

وذلك بإضافة وطرح \hat{y}_i ($\hat{y}_i = b_0 + b_1 \hat{x}_i$) من الطرف الأيمن وتربيع طرفي المعادلة والتجميع مع ملاحظة أنه يمكن إثبات أن المقدار التالي مساوي للصفر :

$$\Sigma(\hat{y}_i - \bar{y})(y_i - \hat{y}_i)^2 = 0.$$

وعلى ذلك :

$$\Sigma(y_i - \bar{y})^2 = \Sigma(\hat{y}_i - \bar{y})^2 + \Sigma(y_i - \hat{y}_i)^2.$$

أي أن مجموع المربعات حول المتوسط يساوي مجموع مربعات الانحدار مضاف إلى مجموع المربعات حول الانحدار ويمكن الرمز إليه بالتساوية التالية :

$$SSTO = SSR + SSE.$$

وهناك صيغ مبسطة للقيم $SSTO$, SSR , SSE حيث :

$$SSTO = SYY = \Sigma y_i^2 - \frac{\Sigma y_i^2}{n},$$

$$SSR = \frac{(SXY)^2}{SXX},$$

$$SSE = SSTO - SSR.$$

من الناحية الإحصائية نجد أن لكل مجموع مربعات درجات حرية خاصة به ، فإذا كان لدينا n من المشاهدات فإن توزيع درجات الحرية يكون على الشكل الموضح في جدول (٣-١٠)

جدول (٣-١٠)

درجات الحرية	مجموع المربعات
1	مجموع مربعات الانحدار
n-2	مجموع مربعات الخطأ
n-1	مجموع المربعات الكلية

بقسمة مجموع المربعات بـ درجات الحرية الخاصة به نحصل على ما يسمى متوسط المربعات

mean squares ويعتبر تباين العينة s^2 مثال لمتوسط المربعات. وعلى ذلك متوسط مجموع

مربعات الانحدار نرمز له بالرمز **MSR** ، هو :

$$MSR = \frac{SSR}{1}.$$

ومتوسط مربعات الخطأ ، نرمز له بالرمز **MSE** ، هو :

$$MSE = \frac{SSE}{n-2}.$$

من النتائج السابقة يمكن اشتقاق جدول تحليل التباين ، ANALYSIS OF VARIANCE ، للاختصار جدول ANOVA ، والموضح في جدول (١٠-٤) . الآن :

لاختبار فرض العدم $H_0 : \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ وباعتبار أن فرض العدم صحيح فإن :

$$f = \frac{MSR}{MSE},$$

قيمة لمغير عشوائي F يتبع توزيع F بدرجات حرية $v_1 = 1, v_2 = n - 2$ لمستوى معنوية α منطقة الرفض $F > f_{\alpha}(1, n - 2)$ حيث $f_{\alpha}(1, n - 2)$ تستخرج من جدول توزيع F في ملحق (٦) أو ملحق (٧) بدرجات حرية $v_1 = 1, v_2 = n - 2$. إذا وقعت f في منطقة الرفض نرفض H_0 .

جدول (١٠-٤)

الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الاتحادار	1	SSR	$MSR = \frac{SSR}{1}$
الخطأ	n-2	SSE	$MSE = \frac{SSE}{n-2}$
الكلي	n-1	SSTO	

مثال (١٠-٣) لزوج القياسات المعطاة في جدول (١٠-٥) :

المطلوب : (أ) إيجاد معادلة الانحدار الخطي المقدرة .

(ب) اختبار فرض العدم $H_0 : \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ عند

مستوى معنوية $\alpha = 0.05$.

جدول (١٠-٥)

x	4	6	2	5	7	6	3	8	5	3	1	5
y	197	272	100	228	327	279	148	377	238	142	66	239

الحل .

$$n=12 \quad \Sigma x_i = 55 \quad \Sigma x_i y_i = 14060$$

$$\Sigma y_i = 2613, \quad \bar{x} = 4.58333, \quad \bar{y} = 217.75,$$

$$\Sigma x_i^2 = 299, \quad \Sigma y_i^2 = 661865,$$

$$SXY = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

$$= 14060 - \frac{(55)(2613)}{12} = 2083.75,$$

$$SXX = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

$$= 299 - \frac{(55)^2}{12} = 46.91667,$$

$$b_1 = \frac{SXY}{SXX} = \frac{2083.75}{46.91667} = 44.41385,$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

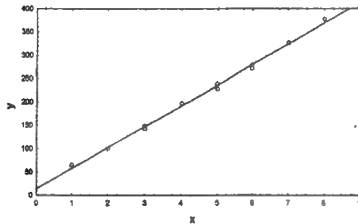
$$= 217.75 - (44.41385)(4.58333)$$

$$= 14.187.$$

وعلى ذلك فإن معادلة خط الانحدار المقترنة هي :

$$\hat{y}_x = 14.187 + 44.41385x.$$

والموضحة في شكل (٥-١٠) .



شكل (٥-١٠)

الآن نحسب :

$$SSTO = SYT = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

$$= 661865 - \frac{(2613)^2}{12} = 92884.25,$$

$$SSR = \frac{(SXY)^2}{SXX} = \frac{(2083.75)^2}{46.91667} = 92547.362.$$

ويطرح SSR من SSTO نحصل على :

$$\begin{aligned} SSE &= SSTO - SSR \\ &= 92884.25 - 92547.362 \\ &= 336.888, \end{aligned}$$

$$MSR = \frac{SSR}{1} = \frac{92547.362}{1} = 92547.362.$$

$$MSE = \frac{SSE}{n-2} = \frac{336.888}{10} = 33.6888.$$

جدول تحليل التباين معطى في جدول (٩-١٠).

جدول (٩-١٠)

متوسط مجموع المربعات	مجموع المربعات	درجات الحرية	مصدر الاختلاف
92547.362	92547.362	1	الانحدار
33.6888	336.888	10	الخطأ
	92884.25	11	الكلية

$$f = \frac{MSR}{MSE} = \frac{92547.362}{33.6888} = 2747.126.$$

$f_{0.05}(1,10) = 4.96$ والمستخرجة من جدول توزيع F في ملحق (٦) بدرجات حرية

$v_1 = 1, v_2 = 10$. منطقة الرفض $F > 4.96$. وبما أن f تقع في منطقة الرفض نرفض H_0 .

(٩-٢-١٠) تقدير σ^2 Estimating σ^2

في الحقيقة ، يفضل الحصول على تقدير للمعلمة σ^2 لا يعتمد على النموذج . عموماً ، مثل هذا التقدير يمكن الحصول عليها فقط في ثنائيات البيانات التي تحتوي على قيم متعددة لـ y عند كل قيمة من قيم x أو عند توفر بعض المعلومات القليلة . في حالة عدم توفر هذه الحسابات فإن التقدير للمعلمة σ^2 يعتمد على النموذج ، أي يكون دالة في مجموع مربعات الأخطاء SSE ، وبحسب من المعادلة الآتية :

$$s^2 = \frac{SSE}{n-2} = MSE$$

أي أن s^2 يساوي متوسط مجموع مربعات الخطأ وهو تقدير غير متحيز للمعلمة σ^2 . التقدير بنقطة للمعلمة σ هو $\sqrt{s^2}$ والذي يسمى الخطأ المعياري للاختبار **standard error of regression**. للمثال (٣-١٠) ومن جدول (٦-١٠) فإن :

$$s = \sqrt{MSE} = \sqrt{33.6888} = 5.804.$$

Coefficient of Simple Determination (٦-٢-١٠) معامل التحديد البسيط

يعرف معامل التحديد البسيط r^2 كالآتي :

$$r^2 = \frac{SSR}{SSTO} = \frac{SSTO - SSE}{SSTO} = 1 - \frac{SSE}{SSTO}$$

للمثال (٣-١٠) ومن جدول (٦-١٠) فإن :

$$r^2 = \frac{92547.362}{92884.25} = 0.996.$$

يأخذ r^2 الواحد الصحيح عندما تقع القيم y_1, y_2, \dots, y_n على خط الانحدار المقدر وعلى ذلك فإن $SSE = 0$. وذلك لأن :

$$SSR = SSTO - SSE$$

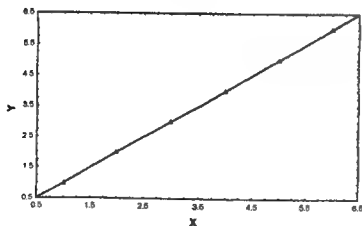
وفي هذه الحالة فإن :

$$SSR = SSTO - 0.0 = SSTO$$

وعلى ذلك فإن :

$$r^2 = \frac{SSR}{SSTO} = \frac{SSTO}{SSTO} = 1$$

هذه الحالة موضحة في شكل (٦-١٠) .

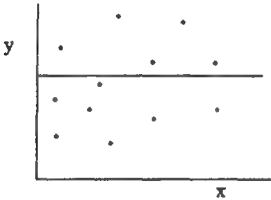


شكل (٦-١٠)

عندما $r^2 = 0$ فهذا يدل على عدم وجود علاقة خطية بين المتغيرين وبالتالي فإن $SSR = 0$ ومنها :

$$r^2 = \frac{SSR}{SSTO} = \frac{0.0}{SSTO} = 0.0$$

في هذه الحالة فإن معادلة الانحدار المقدرة سوف تكون موازية للمحور الأفقي ، أي أن $b_1 = 0$ كما هو موضح في شكل (٧-١٠)

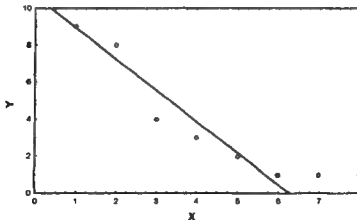


شكل (٧-١٠)

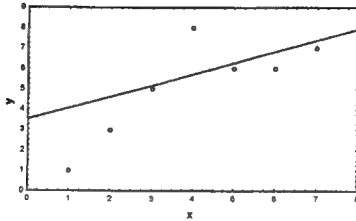
معامل التحديد دائما موجب وتتراوح قيمته بين الصفر والواحد الصحيح أي أن :

$$0 \leq r^2 \leq 1.$$

يتضح من شكل (٨-١٠) ، حيث $r^2 = -0.9025$ ، أن المشاهدات تقرب بدرجة كبيرة من خط الانحدار المقدر وذلك بالمقارنة للمشاهدات في شكل (٩-١٠) حيث $r^2 = 0.3249$.



شكل (٨-١٠)



شكل (٩-١٠)

Estimation of the parameters β_0, β_1 تقدير المعلمتين (٧-٢-١٠)

في دراستنا السابقة في البند (٧-١٠) عرفنا أن b_0, b_1 تقديرين للمعلمتين الحقيقيتين β_0, β_1 وذلك بالاعتماد على عينة عشوائية من الحجم n . بتكرار المعاينة من الحجم n وحساب b_0, b_1 لكل عينة فإن القيمتين b_0, b_1 سوف يختلفان من عينة إلى أخرى. وعلى ذلك فإن التقديرين b_0, b_1 قيمتين لتغيرين عشوائيتين B_0, B_1 . وبما أن قيمة x لازالت ثابتة، فإن قيم β_0, β_1 سوف تعتمد على قيم y أو بدقة أكثر، على قيم المتغيرات العشوائية Y_1, Y_2, \dots, Y_n المستقلة والتي تتبع توزيعاً طبيعياً، وباستخدام نظريات الإحصاء يمكن إثبات أن المتغير العشوائي B_1 أيضاً يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{B_1} = \beta_1$$

وتباين :

$$\sigma_{B_1}^2 = \frac{\sigma^2}{SXX}.$$

وتبعاً لنظرية (٨-١) فإن :

$$Z = \frac{B_1 - \beta_1}{\sqrt{\frac{\sigma^2}{SXX}}}$$

متغير عشوائي يتبع التوزيع الطبيعي القياسي .

عادة الانحراف المعياري ، σ ، مجهول ويستبدل بالإحصاء S في صيغة Z لنحصل على

$$T = \frac{B_1 - \beta_1}{\sqrt{\frac{S^2}{SXX}}}$$

حيث T متغير عشوائي له توزيع t بدرجات حرية $n-2$. سوف نستخدم المتغير T في إيجاد $100(1-\alpha)\%$ فترة ثقة للمعلمة β_1 حيث :

$$P(-t_{\alpha/2} < \frac{B_1 - \beta_1}{\sqrt{S^2/SXX}} < t_{\alpha/2}) = 1 - \alpha.$$

ويأتباع الخطوات الجبرية التي استخدمناها في الفصل الثامن يمكن كتابة :

$$P(B_1 - t_{\alpha/2} \sqrt{\frac{S^2}{SXX}} < \beta_1 < B_1 + t_{\alpha/2} \sqrt{\frac{S^2}{SXX}}) = 1 - \alpha.$$

لعينة عشوائية معطاة من الحجم n نحسب s و SXX وذلك للحصول على $100(1-\alpha)\%$ فترة ثقة كالتالي :

$$b_1 - t_{\alpha/2} \sqrt{\frac{s^2}{SXX}} < \beta_1 < b_1 + t_{\alpha/2} \sqrt{\frac{s^2}{SXX}}.$$

مثال (١٠-٤) أوجد 95% فترة ثقة للمعلمة β_1 في معادلة الانحدار $\mu_{y|x} = \beta_0 + \beta_1 x$ بالاعتماد على البيانات في جدول (١٠-٥) .

الحل . للمثال (١٠-٣) فإن :

$$SXX = 46.91667, \quad b_1 = 44.41385, \quad s^2 = 33.6888$$

باستخدام جدول توزيع t في ملحق (٤) فإن $t_{0.025} = 2.228$ بدرجات حرية $v = 10$. وعلى ذلك فإن 95% فترة ثقة للمعلمة β_1 نحسب كالآتي :

$$44.41385 - 2.228 \sqrt{\frac{33.6888}{46.91667}} < \beta_1 < 44.41385 + 2.228 \sqrt{\frac{33.6888}{46.91667}}$$

أي أن :

$$44.41385 - 2.228(0.847382) < \beta_1 < 44.41385 + 2.228(0.847382)$$

والتي تختصر إلى :

$$42.5 < \beta_1 < 46.3.$$

لاختبار فرض المدم $H_0: \beta_1 = \beta_1^*$ ضد فرض بديل مناسب يمكننا استخدام توزيع

t بدرجات حرية $n-2$ للحصول على منطقة الرفض. قررنا سوف يعتمد على القيمة :

$$t = \frac{b_1 - \beta_1^*}{\sqrt{s^2 / SXX}}$$

مثال (١٠-٥) باستخدام القيمة المقدرة $b_0=44.41385$ في مثال (١٠-٣) ، اختبار فرض العدم أن $\beta_1 = 0$ عند مستوى معنوية $\alpha = 0.05$.
الحل .

$$H_0 : \beta_1 = 0,$$

$$H_1 : \beta_1 \neq 0.$$

$$\alpha = 0.05.$$

$$T < -2.228 \text{ أو } T > 2.228 \text{ ومنطقة الرفض } t_{0.025}=2.228$$

$$t = \frac{b_1 - 0.0}{\sqrt{s^2 / SXX}}$$

$$= \frac{44.41385}{\sqrt{\frac{33.6888}{46.91667}}} = \frac{44.41385}{0.847382} = 52.41.$$

وبما أن t تقع في منطقة الرفض لرفض H_0 .

أيضا المتغير العشوائي B_0 له توزيع طبيعي بمتوسط :

$$\mu_{B_0} = \beta_0$$

وتباين :

$$\sigma_{B_0}^2 = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{SXX} \right)$$

وتبعا لنظرية (٨-١) فإن :

$$Z = \frac{B_0 - \beta_0}{\sqrt{\sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{SXX} \right)}}$$

متغير عشوائي يتبع التوزيع الطبيعي القياسي وباستبدال σ بالإحصاء S في صيغة Z نحصل على المتغير العشوائي :

$$T = \frac{B_0 - \beta_0}{\sqrt{S^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{SXX} \right)}}$$

والذي يتبع توزيع t بدرجات حرية $n - 2$.

سوف يستخدم المظهر T للحصول على $100(1 - \alpha)\%$ فترة ثقة للمعلمة β_0 كالتالي :

$$b_0 - t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{SXX} \right)} < \beta_0 < b_0 + t_{\alpha/2} \sqrt{s^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{SXX} \right)}.$$

مثال (١٠-٦) أوجد 95% فترة ثقة للمعلمة β_0 في خط الانحدار $\mu_{y|x} = \beta_0 + \beta_1 x$

بالاعتماد على البيانات في جدول (١٠-٣) .

الحل . من المثال (١٠-٣) فإن :

$$s^2 = 33.6888 , SXX = 46.91667 , \bar{x} = 4.58333 , b_0 = 14.187$$

وعلى ذلك 95% فترة ثقة للمعلمة β_0 تعطى على الشكل :

$$14.187 - 2.228 \sqrt{33.6888 \left[\frac{1}{12} + \frac{4.58333^2}{46.91667} \right]} < \beta_0$$

$$14.187 + 2.228 \sqrt{33.6888 \left[\frac{1}{12} + \frac{4.58333^2}{46.91667} \right]}.$$

أي أن :

$$14.187 - 2.228(4.2298) < \beta_0 < 14.187 + 2.228(4.2298)$$

والتي تختزل إلى :

$$4.8 < \beta_0 < 23.6.$$

لاختبار فرض العدم $H_0 : \beta_0 = \beta_0^*$ ضد أي فرض مناسب فإننا مرة أخرى سوف

نستخدم توزيع t بدرجات حرية $n - 2$ للحصول على منطقة الرفض وبالتالي فإن قرارنا

سوف يعتمد على القيمة :

$$t = \frac{b_0 - \beta_0^*}{\sqrt{s^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{SXX} \right)}}$$

الطريقة المتبعة موضحة في المثال التالي :

مثال (٧-١٠) باستخدام القيمة المقدرة $b_0 = 14.187$ في مثال (٣-١٠) ، اختبار فرض العلم $H_0 : \beta_0 = 0$ ضد الفرض البديل $H_0 : \beta_0 \neq 0$ عند مستوى معنوية $\alpha = 0.05$.
الحل .

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

مستوى معنوية $\alpha = 0.05$

$$t = \frac{b_0 - 0.0}{\sqrt{s^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{SXX} \right)}}$$

$$= \frac{14.187}{\sqrt{33.6888 \left[\frac{1}{12} + \frac{4.58333^2}{46.91667} \right]}} = 3.354.$$

$t_{0.025} = 2.228$ ومنطقة الرفض $T > 2.228$ أو $T < -2.228$ وحيث أن t وقعت في منطقة الرفض نرفض H_0 .

(٨-٢-١٠) التنبؤ Prediction

يمكن استخدام المعادلة $\hat{y}_x = b_0 + b_1x$ للتنبؤ بقيمة $\mu_{Y|x}$ ، حيث x' ليس من الضروري أن تكون واحدة من x_1, x_2, \dots, x_n في العينة العشوائية من الحجم n للملاحظات $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. أيضا يمكن استخدام المعادلة $\hat{y}_x = b_0 + b_1x$ للتنبؤ بقيمة واحدة y_x للمتغير Y . سوف نتوقع أن خطأ التنبؤ سوف يكون أعلى في حالة قيمة واحدة متباً بما عنه في حالة التنبؤ بالمتوسط وهذا سوف يؤثر على طول فترة الثقة للمعالم المراد تقديرها .

عند الرغبة في الحصول على فترة ثقة للمعلمة $\mu_{Y|x}$ يكون من الضروري إتساع التوزيع العيني للفروق بين القيم \hat{y}_x والتي نحصل عليها من خط الانحدار المقدر بتكرار المعاينة والقيمة الحقيقية المقابلة $\mu_{Y|x}$ من خط الانحدار الحقيقي . باستخدام نظريات الإحصاء يمكن إثبات أن التوزيع العيني للإحصاء $\hat{Y}_x - \mu_{Y|x}$ يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{Y}_x - \mu_{Y|x}} = E[\hat{Y}_x - \mu_{Y|x}] = 0$$

وتباين :

$$\sigma_{\hat{Y}_{x'} - \mu_{Y|x'}}^2 = \sigma^2 \left[\frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX} \right].$$

في التطبيق يستبدل σ^2 بالقيمة s^2 والتي تمثل قيمة للإحصاء S^2 وعلى ذلك ، فإن الإحصاء :

$$T = \frac{\hat{Y}_{x'} - \mu_{Y|x'}}{\sqrt{S^2 \left[\frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX} \right]}}$$

له توزيع t بدرجات حرية $n-2$. يمكن الحصول على $100(1-\alpha)\%$ فترة ثقة للمعلمة $\mu_{Y|x'}$

من الصيغة التالية :

$$\hat{Y}_{x'} - t_{\alpha/2} \sqrt{s^2 \left[\frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX} \right]} < \mu_{Y|x'} < \hat{Y}_{x'} + t_{\alpha/2} \sqrt{s^2 \left[\frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX} \right]}.$$

مثال (١٠-٨) باستخدام البيانات في جدول (١٠-٥) أوجد 95% فترة ثقة للمعلمة

$$\mu_{Y|4}$$

الحل . من معادلة الانحدار المقدرة فإن :

$$\begin{aligned}\hat{y}_4 &= 14.187 + (44.41385)(4) \\ &= 191.84.\end{aligned}$$

عرفنا مما سبق أن :

$$SXX = 46.91667, \bar{x} = 4.58333, s^2 = 33.6888,$$

$t_{0.025} = 2.228$ بدرجات حرية 10 . وعلى ذلك 95% فترة ثقة للمعلمة $\mu_{Y|4}$ هي :

$$\begin{aligned}191.84 - 2.228 \sqrt{33.6888 \left[\frac{1}{12} + \frac{(4 - 4.58333)^2}{46.91667} \right]} &< \mu_{Y|4} < \\ 191.84 + 2.228 \sqrt{33.6888 \left[\frac{1}{12} + \frac{(4 - 4.58333)^2}{46.91667} \right]} &.\end{aligned}$$

أي أن :

$$191.84 - (2.228)(1.7469) < \mu_{Y|4} < 191.84 + (2.228)(1.7469)$$

والتي تختصر إلى :

$$187.94791 < \mu_{Y|4} < 195.73209.$$

للحصول على $100\%(1-\alpha)$ فترة ثقة لأي قيمة مفردة $y_{x'}$ للمتغير $Y | x'$ ، يكون من الضروري تقدير التباين للفروق بين القيم $\hat{y}_{x'}$ المحسوبة من خط الانحدار المقدر بتكرار المعاينة والقيمة المقابلة الحقيقية $y_{x'}$. باستخدام نظريات الإحصاء يمكن إثبات أن التوزيع العيني للإحصاء $\hat{y}_{x'} - y_{x'}$ يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{y}_{x'} - y_{x'}} = 0$$

ولبيان :

$$\sigma_{\hat{y}_{x'} - y_{x'}}^2 = \sigma^2 \left[1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX} \right]$$

في التطبيق يستبدل σ^2 بالقيمة s^2 والتي تمثل قيمة للإحصاء S^2 وعلى ذلك ، فإن

الإحصاء :

$$T = \frac{\hat{y}_{x'} - y_{x'}}{\sqrt{S^2 \left[1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX} \right]}}$$

يتبع توزيع t بدرجات حرية $n-2$.

يمكن الحصول على $100\%(1-\alpha)$ فترة ثقة مفردة $y_{x'}$ من الصيغة التالية :

$$\hat{y}_{x'} - t_{\alpha/2} \sqrt{1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX}} < y_{x'} < \hat{y}_{x'} + t_{\alpha/2} \sqrt{1 + \frac{1}{n} + \frac{(x' - \bar{x})^2}{SXX}}$$

مثال (٩-١٠) . للبيانات في جدول (٩-١٠) أوجد (أ) 95% فترة ثقة لـ y_4 (ب) وضع بيانيا 95% فترة ثقة لـ $y_{x'}$.

الحل . (أ) $\bar{x} = 4.58333$ ، $s^2 = 33.6888$ ، $n = 12$ وعلى ذلك 95% فترة ثقة

لـ y_4 هي :

$$191.84 - 2.228 \sqrt{33.6888 \left[1 + \frac{1}{12} + \frac{(4 - 4.58333)^2}{46.91667} \right]} < y_4 < 191.84 + 2.228 \sqrt{33.6888 \left[1 + \frac{1}{12} + \frac{(4 - 4.58333)^2}{46.91667} \right]}$$

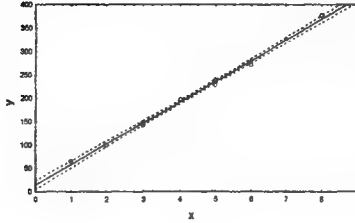
أي أن :

$$191.84 - 2.228(6.061) < y_4 < 191.84 + 2.228(6.06)$$

والتي تختصر إلى :

$$178.3 < y_4 < 205.3$$

(ب) يوضح شكل (١٠-١٠) 95% فترة ثقة لـ y_x .



شكل (١٠-١٠)

(٩-٢-١٠) اختبار خطية الانحدار Test for Linearity of Regression

لمشكلة معطاة فإننا نفترض إما أن الانحدار خطي ونتبع خطوات التقدير التي تناولناها في البند (٣-٢-١٠) أو نستنتج أن الانحدار غير خطي وفي هذه الحالة سوف نتبع الخطوات التي سوف نتناولها في البند (٣-١٠). في الجزء التالي سوف نتناول اختبار الخطية في حالتين.

الحالة الأولى عندما تكون σ^2 معلومة (التباين معلوم)

لاختبار فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_1 : النموذج غير خطي فإن:

$$\chi^2 = \frac{SSE}{\sigma^2}$$

قيمة لمعير عشوائي X^2 له توزيع χ^2 بدرجات حرية $v = n - 2$ وذلك بالفرض صحة فرض العدم. لمستوى معنوية α فإن منطقة الرفض $\chi^2 > \chi^2_{\alpha}$ حيث χ^2_{α} نستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $v = n - 2$. إذا وقعت χ^2 في منطقة الرفض لرفض H_0 . مثال (١٠-١٠) لأزواج القياسات في جدول (٧-١٠) اختبار فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_1 : النموذج غير خطي وذلك عند مستوى معنوية $\alpha = 0.01$ ونخت فرض أن $\sigma^2 = 1$.

جدول (٧-١٠)

x	0.345	0.287	0.251	0.225	0.207	0.186	0.161	0.132	0.084	0.060
y	367	311	295	268	253	239	220	213	193	192

الحل . لاختبار فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_1 : النموذج غير خطي نحسب جدول تحليل التباين والمعطى في جدول (٨-١٠) ونحسب القيمة :

$$\chi^2 = \frac{SSE}{\sigma^2} = \frac{21.95}{1} = 21.95.$$

بمسوي معنوية $\alpha = 0.01$ فإن $\chi^2_{0.01} = 20.09$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $v = n - 2 = 10 - 2 = 8$. منطقة الرفض $\chi^2 > 20.09$. وبما أن χ^2 تقع في منطقة الرفض نرفض H_0 ، أي أن العلاقة بين المتغيرين لا تتبع النموذج الخطي ويجب أن نبحث عن نموذج آخر.

جدول (٨-١٠)

متوسط مجموع المربعات	مجموع المربعات	درجات الحرية	مصدر الاختلاف
342.0	342.0	1	الانحدار
2.744	21.95	8	الخطأ
	363.95	9	الكلية

الحالة التالية عندما تكون σ^2 مجهولة :

لاختبار فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_1 : النموذج غير خطي نختار عينة عشوائية من الحجم n من المشاهدات بحيث أن لكل قيمة من x يوجد أكثر من قيمة لـ y . أي أن العينة التي حجمها n تحتوي على قيم مختلفة من x عددها k بحيث تحتوي العينة على n_1 من القيم المشاهدة $y_{11}, y_{12}, \dots, y_{1n_1}$ المقابلة للقيمة x_1 ، n_2 من القيم المشاهدة $y_{21}, y_{22}, \dots, y_{2n_2}$ المقابلة للقيمة x_2 و... و n_k من القيم المشاهدة

$$n = \sum_{i=1}^k n_i \quad \text{حيث } x_k \text{ المقابلة للقيمة } y_{k1}, y_{k2}, \dots, y_{kn_k}.$$

لإجراء الاختبار نتبع الخطوات التالية :

(أ) نحسب مجموع مربعات الخطأ الخالص من x_1 pure error of sum squares

من الصيغة التالية :

$$\sum_{u=1}^{n_1} (y_{1u} - \bar{y}_1)^2$$

، حيث $\bar{y}_i = \sum_{u=1}^{n_i} y_{iu} / n_i$ ، بدرجات حرية $n_i - 1$. بنفس الشكل يمكن حساب مجموع مربعات الخطأ الخالص من x_2 و ... و x_k . مجموع مربعات الخطأ الخالص الكلي يحسب من الصيغة التالية :

$$SSP = \sum_{i=1}^k \sum_{u=1}^{n_i} (y_{iu} - \bar{y}_i)^2$$

بدرجات حرية :

$$n_e = \sum_{i=1}^k (n_i - 1) = \sum_{i=1}^k n_i - k$$

(ب) نحسب متوسط مجموع مربعات الخطأ الخالص من الصيغة التالية :

$$s_e^2 = \frac{SSP}{n_e}$$

والذي يعتبر تقدير للمعلمة σ^2

(ج) نحسب مجموع مربعات قصور جودة التوفيق sum squares lake of fit كالتالي :

$$SSL = SSE - SSP$$

بدرجات حرية $n_L = (n-2) - n_e$.

من الحسابات السابقة فإن جدول تحليل التباين يكون كما هو موضح في جدول (٩-١٠)

جدول (٩-١٠)

متوسط مجموع المربعات	مجموع المربعات	درجات الحرية	مصدر الاختلاف
SSR/1 SSE/n-2	SSR SSE	1 n-2	الانحدار الخطأ
		n-1	الكلي
MSL=SSL/n _L s _e ² = SSP / n _e	SSL SSP	n _L n _e	قصور جودة التوفيق الخطأ الخالص

نحسب قيمة f' كالتالي :

$$f' = \frac{MSL}{s_e^2} .$$

بافتراض صحة فرض العدم فإن f' تمثل قيمة لمغير عشوائي F له توزيع F بدرجات حرية n_L, n_e . عند مسري معنوية α نحسب منطقة الرفض كالتالي : $F > f_{\alpha}(n_L, n_e)$ حيث :

$f_{\alpha}(n_L, n_e)$ تستخرج من جدول توزيع F في ملحق (٦) أو ملحق (٧) بدرجات حرية $v_1 = n_L, v_2 = n_e$. إذا وقعت f' في منطقة الرفض نرفض H_0 .

مثال (١٠-١١) لارواج القياسات في جدول (١٠-١٠) اختبر فرض العدم H_0 : النموذج خطي ضد الفرض البديل H_0 : النموذج غير خطي عند مستوى معنوية $\alpha = 0.05$.

جدول (١٠-١٠)

الملاحظات	x	y	الملاحظات	x	y	الملاحظات	x	y
1	1.3	2.3	9	3.7	1.7	17	5.3	3.5
2	1.3	1.8	10	4.0	2.8	18	5.3	2.8
3	2.0	2.8	11	4.0	2.8	19	5.3	2.1
4	2.0	1.5	12	4.0	2.2	20	5.7	3.4
5	2.7	2.2	13	4.7	5.4	21	6.0	3.2
6	3.3	3.8	14	4.7	3.2	22	6.0	3.0
7	3.3	1.8	15	4.7	1.9	23	6.3	3.0
8	3.7	3.7	16	5.0	1.8	24	6.7	5.9

الحل .

H_0 : النموذج الخطي :

H_1 : النموذج غير خطي :

$\alpha = 0.05$

سوف نوجد مجموع مربعات الخالص ثم مجموع مربعات قصور التوفيق كالتالي :

مجموع مربعات الخطأ الخالص عند $x = 1.3$ هو :

$$(2.3)^2 + (1.8)^2 - \{ (2.3 + 1.8)^2 / 2 \} = 0.125$$

بدرجة حرية واحدة ($n_1 = 2 - 1 = 1$) .

مجموع مربعات الخطأ الخالص عند $x = 2.0$ هو :

$$(2.8)^2 + (1.5)^2 - \{ (2.8 + 1.5)^2 / 2 \} = 0.845$$

بدرجة حرية واحدة ($n_2 = 2 - 1 = 1$) . بنفس الطريقة يتم حساب مجموع مربعات الخطأ

الخالص للقيم الباقية من x كما في جدول (١٠-١١) .

جدول (١٠-١١)

مستوى x	$\Sigma(y_{iu} - \bar{y}_i)^2$	درجات حرية
1.3	0.125	1
2.0	0.845	1
3.3	2.000	1
3.7	2.000	1

4.0	0.240	2
4.7	6.260	2
5.3	0.980	2
6.0	0.020	1
المجموع	12.470	11

جدول تحليل التباين معطى في جدول (١٠-١٢).

جدول (١٠-١٢)

f المعسوبة	متوسط مجموع المربعات	مجموع المربعات	درجات الحرية	مصدر الاختلاف
$f = \frac{6.326}{0.963}$ $= 6.569$ عند مستوى $\alpha=0.05$ معنوية	6.326	6.326	1	الانحدار
	$0.963=s^2$	21.192	22	الخطأ
$f' = \frac{0.793}{1.134}$ $= 0.699$	0.793=MSL $1.134=s_e^2$	8.722 12.470	11 11	قصور التوفيق الخطأ الخالص

من جدول (١٠-١٢) فإن القيمة $f' = 0.699$ غير معنوية لأنها أقل من الواحد.

(١٠-٣) بعض نماذج الانحدار الغير خطية

Some Nonlinear Regression Models

يوجد العديد من النماذج الرياضية الغير خطية المستخدمة في تحليل العلاقات الاقتصادية والاجتماعية ... الخ. سوف تقتصر دراستنا في هذا الجزء على النموذج الأسى ونموذج القوى.

The Exponential Model (١٠-٣-١) النموذج الأسى

معادلة النموذج الأسى تكون على الصورة التالية :

$$\mu_{Y|X} = \gamma \delta^x$$

حيث γ, δ ثابتان والمطلوب تقديرهما من البيانات بالتقديرين d, c على التوالي . يمكن تقدير

$\mu_{Y|X}$ بالقيمة \hat{y}_x من منحنى الانحدار المقدر التالي :

$$\hat{y}_x = c d^x.$$

بأخذ لوغاريتمات الطرفين (للأساس e) في المعادلة السابقة فإن منحنى الانحدار المقدر يمكن

كتابته على الشكل :

$$\ln \hat{y}_x = \ln c + (\ln d)x,$$

وكل زوج من المشاهدات في العينة يحقق العلاقة :

$$\ln y_i = \ln c + (\ln d) x_i + e_i = b_0 + b_1 x_i + e_i,$$

حيث أن : $b_1 = (\ln d)$, $b_0 = \ln c$. وعلى ذلك يمكن إيجاد b_0 , b_1 بالصيغ المستخدمة

لنموذج الانحدار الخطي ، التي سبق أن تناولناها ، باستخدام النقاط $(x_i, \ln y_i)$ ثم إيجاد c, d

بأخذ القيم المقابلة للوغاريتمات لـ b_0, b_1 على التوالي، أي أن :

$$d = \exp(b_1), c = \exp(b_0).$$

طريقة المربعات الصغرى لتوفيق منحنى أسى لثقة من المشاهدات موضحة في المثال التالي .

مثال (١٢-١٠) : لزوج القياسات في جدول (١٣-١٠) أوجد معادلة الانحدار المقدرة تحت فرض النموذج الأسى .

جول (١٣-١٠)

x	1	2	3	4	5	6	7
y	304	341	393	457	548	670	882

الحل . بوضع $y'_i = \ln y_i$ فإن : $\Sigma y'_i = 43.243148$

$$n = 7, \quad \Sigma x_i = 28, \quad \Sigma x_i^2 = 140$$

$$\Sigma x_i y'_i = 177.85134, \quad \bar{x} = 4, \quad \frac{\Sigma y'_i}{n} = 6.1775926$$

$$b_1 = \frac{177.85134 - \frac{(28)(43.243148)}{7}}{140 - \frac{(28)^2}{7}}$$

$$= 0.174241.$$

$$b_0 = 6.1775926 - (0.174241)(4)$$

$$= 5.4806286.$$

معادلة الانحدار المقدرة هي :

$$\hat{y}_x = 5.4806286 + 0.174241x.$$

وعلى ذلك :

$$\ln d = b_1 = 0.174241, \quad \ln c = b_0 = 5.4806286,$$

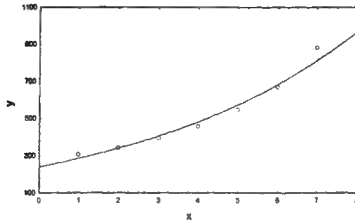
$$d = \exp(b_1) = 1.1903424, c = \exp(b_0) = 239.99752.$$

وبالتالي فإن منحنى الانحدار المقدر بالمربعات الصغرى هو :

$$\hat{y}_x = c d^x$$

$$= (239.99752)(1.1903424)^x$$

والتمثيل البياني لها موضح في شكل (١٠-١١) .



شكل (١٠-١١)

Power Model (١٠-٣-٢) نموذج القوى

معادلة نموذج القوى تكون على الصورة التالية :

$$\mu_{y|x} = a_0 x^{a'_0}$$

حيث a_0, a'_0 ثابتان والمطلوب تقديرهما من البيانات بالتقديرين d_0, c_0 على التوالي. يمكن

تقدير $\mu_{y|x}$ بالقيمة \hat{y}_x من منحنى الانحدار المقدر التالي :

$$\hat{y}_x = c_0 x^{d_0}.$$

بأخذ لوغاريتمات الطرفين (للأساس e) فإن منحنى الانحدار يمكن كتابته على الشكل :

$$\ln \hat{y}_x = \ln c_0 + d_0 (\ln x)$$

كل زوج من المشاهدات في العينة يحقق العلاقة :

$$\ln y_i = \ln c_0 + d_0 (\ln x_i) + e_i$$

$$= b_0 + b_1 (\ln x_i) + e_i$$

حيث $b_1 = d_0, b_0 = \ln c_0$. وعلى ذلك يمكن إيجاد b_0, b_1 بالصيغ المستخدمة لنموذج

الانحدار الخطي، التي سبق أن تناولناها ، باستخدام النقاط $(\ln x_i, \ln y_i)$ ثم إيجاد c_0, d_0 حيث

$$b_1 = d_0, \ln c_0 = b_0$$

مثال (١٠-١٣) لزوج القياسات في جدول (١٠-١٤) أوجد معادلة الانحدار المقطرة تحت

لرض نموذج القوى .

جدول (١٠ - ١٤)

x	600	600	600	600	500	500	500	500	400	400	400	400
y	2.35	2.65	3.0	3.6	6.4	7.8	9.8	16.5	21.5	24.5	26.0	33.0

الحل .

$$n = 12 , \Sigma \ln x_i = 74.412 , \Sigma \ln y_i = 26.22601 ,$$

$$\Sigma \ln x_i^2 = 461.75874 , \Sigma (\ln x_i)(\ln y_i) = 160.84601 ,$$

$$\Sigma \ln y_i^2 = 67.74609.$$

الحل.

$$b_1 = \frac{160.84601 - \frac{(74.412)(26.22601)}{12}}{461.75874 - \frac{(74.412)^2}{12}}$$

$$= - 5.3996,$$

$$b_0 = \frac{26.22601 - (-5.3996)(74.412)}{12}$$

$$= 35.6684.$$

معادلة الانحدار المقبرة هي :

$$\hat{y}_x = 35.6684 - 5.3996x.$$

وعلى ذلك :

$$\ln c_0 = b_0 = 35.6684$$

أي أن :

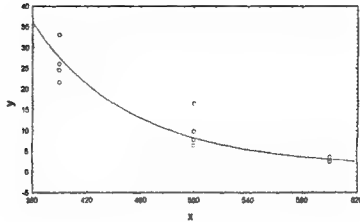
$$c_0 = \exp(b_0) = 3.094491530 \cdot 10^{15}$$

$$d_0 = b_1 = -5.3996$$

والمعادلة الأساسية المقبرة هي :

$$\hat{y}_x = c_0 x^{d_0} = (3.094491530 \cdot 10^{15}) \cdot x^{-5.3996}.$$

والتمثيل البياني لها موضح في شكل (١٠ - ١٢).



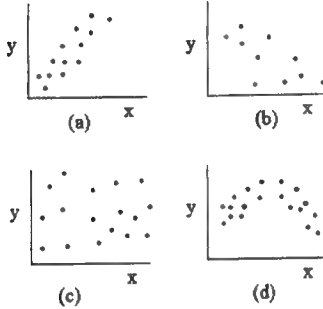
شكل (١٠-١٢)

(١٠-٤) معامل الارتباط الخطي البسيط

The Simple Linear Correlation Coefficient

في مشكلة الانحدار كان اهتمامنا بالتنبؤ بتغير وذلك من المعلومات عن المتغيرات المستقلة ، بينما في مشكلة الارتباط فإن اهتمامنا سوف يكون في قياس العلاقة بين متغيرين أو أكثر. مرة أخرى فإن قيم المتغيرات المستقلة كانت ثابتة في مشكلة الانحدار. الآن سوف يختلف الوضع. سوف نعرف معامل الارتباط الخطي بأنه مقياس للعلاقة بين متغيرين عشوائيين X, Y . وسوف نرمز له بالرمز r . سوف نفترض أن المتغيران X, Y لهما توزيع احتمالي ثنائي. لحساب معامل الارتباط الخطي نختار عينة عشوائية من أزواج المشاهدات (x, y) . إذا كانت نقاط شكل الانتشار تتركز فوق وحول خط انحدار له ميل موجب ، فهذا يدل على ارتباط موجب قوي بين المتغيرين (ارتباط طردي) كما هو موضح في شكل (١٠-١٣) (a). ومن ناحية أخرى ، إذا كانت نقاط شكل الانتشار تتركز فوق وحول خط انحدار له ميل سالب فهذا يدل على ارتباط قوى سالب بين المتغيرين (ارتباط عكسي) كما هو موضح في شكل (١٠-١٣) (b) كلما زاد انتشار نقاط شكل الانتشار حول وفوق خط الانحدار فإن الارتباط يقل عددياً بين المتغيرين ، إذا كانت نقاط شكل الانتشار تنتشر بطريقة عشوائية كما في شكل (١٠-١٣) (c) فهذا يعني أن $r=0$ ونستنتج عدم وجود علاقة بين X, Y . ولما كان معامل الارتباط بين متغيرين يعتبر مقياس للعلاقة الخطية بينهما وعلى ذلك فإن $r=0$ تعني قصور في الخطية ولتست قصور في الارتباط . على سبيل المثال قد تكون هناك علاقة ولكنها علاقة غير خطية . فعلى سبيل المثال إذا

وجدت علاقة قوية من الدرجة الثانية بين المتغيرين X, Y كما هو موضح في شكل (١٠-١٣) (d) فهذا يعني أن $r=0$. يعتبر معامل الارتباط الخطي (معامل بيرسون للارتباط) أو اختصارا معامل الارتباط أكثر مقياسي الارتباط الخطي انتشارا.



شكل (١٠-١٣)

يتم حساب معامل الارتباط من المعادلة التالية :

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left[\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right] \left[\sum y_i^2 - \frac{(\sum y_i)^2}{n} \right]}}$$

$$= \frac{S_{XY}}{\sqrt{S_{XX} S_{YY}}}$$

من الجزء (١٠-٢-٩) يمكن حساب معامل الارتباط كالتالي :

$$r = \pm \sqrt{r^2}$$

حيث r^2 هو معامل التحديد البسيط والإشارة تخص التقدير b_1 . وما أن $0 \leq r^2 \leq 1$ فإن r

يمكن أن تأخذ الإشارة الموجبة أو السالبة ، أي أن :

$$-1 \leq r \leq 1$$

عادة يفضل حساب معامل الارتباط من معادلته وليس من r^2 لصعوبة الحساب.

مثال (١٤-١٠) لدراسة العلاقة بين تركيز الأوزون (X) Ozone (مقاس PPM) وتركيز الكربون (Y) (مقاس $\mu\text{g} / \text{m}^3$) تم الحصول على البيانات المعطاة في جدول (١٥-١٠).

جدول (١٥-١٠)

x	0.066	0.088	0.120	0.050	0.162	0.186	0.057	0.100
y	4.6	11.6	9.5	6.3	13.8	15.4	2.5	11.8
x	0.112	0.055	0.154	0.074	0.111	0.140	0.071	0.110
y	8.0	7.0	20.6	16.6	9.2	17.9	2.8	13.0

أوجد معامل الارتباط البسيط .

الحل .

$$n = 16 , \quad \Sigma x_i = 1.656 , \quad \Sigma y_i = 170.6 ,$$

$$\Sigma x_i^2 = 0.196912 , \quad \Sigma x_i y_i = 20.0397 ,$$

$$\Sigma y_i^2 = 2253.56.$$

$$SXY = \Sigma x_i y_i - \frac{\Sigma x_i \Sigma y_i}{n}$$

$$= 20.0397 - \frac{(1.656)(170.6)}{16}$$

$$= 2.3826,$$

$$SXX = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}$$

$$= 0.196912 - \frac{(1.656)^2}{16} = 0.025516,$$

$$SYY = \Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n} = 2253.56 - \frac{(170.6)^2}{16}$$

$$= 434.5375.$$

وعلى ذلك :

$$r = \frac{SXY}{\sqrt{SXX \cdot SYY}} = \frac{2.3826}{\sqrt{(0.025516)(434.5375)}} = 0.716.$$

في المناقشة السابقة لم نضع فروض قوية على توزيع المجتمع الذي اخبرت منه العينة وذلك للحصول على تقدير بنقطة للمعلمة ρ والتي ترمز إلى معامل ارتباط المجتمع. للحصول على

100%(1-α) فترة ثقة للمعلمة ρ أو اختبارات لفروض تخص ρ فإننا نفترض أن العينة مأخوذة من مجتمع يتبع التوزيع الطبيعي الثنائي the bivariate normal distribution أي أن X, Y متغيرين عشوائيين حيث دالة التوزيع الهامشية لكل من X, Y تتبع التوزيع الطبيعي . يمكن اختبار الاعتدال لقيم X وقيم Y على حدة بطريقة رياضية سوف نتناوباً في البند (١٢-٥) من الفصل الثاني عشر .

اختبارات لفروض وفترات ثقة تخص ρ

Tests Hypotheses and Confidence Intervals Concerning ρ

لاختبار فرض العدم $H_0: \rho = 0$ ضد الفرض البديل $H_1: \rho \neq 0$ أو الفرض البديل $H_1: \rho > 0$ أو الفرض البديل $H_1: \rho < 0$ وبافتراض صحة فرض العدم فإن :

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

هي قيمة لمغير عشوائي T له توزيع t بدرجات حرية $v = n - 2$. وعلى ذلك مستوى معنوية α وللفرض البديل $H_1: \rho \neq 0$ (اختبار ذي جانين) فإن منطقة الرفض سوف تكون $T < -t_{\alpha/2}$ or $T > t_{\alpha/2}$ حيث $t_{\alpha/2}$ هي قيمة t المستخرجة من جدول توزيع t بدرجات حرية $v = n - 2$. للبديل $H_1: \rho > 0$ فإن منطقة الرفض $T > t_{\alpha}$ وللبديل $H_1: \rho < 0$ فإن منطقة الرفض $T < -t_{\alpha}$.

مثال (١٥-١٠) بفرض أن البيانات في مثال (١٤-١٠) مأخوذة من مجتمع يتبع التوزيع الثنائي الطبيعي . المطلوب اختبار فرض العدم $H_0: \rho = 0$ ضد الفرض البديل $H_1: \rho > 0$ وذلك عند مستوى معنوية $\alpha = 0.01$.

الحل .

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.716\sqrt{14}}{\sqrt{1-(0.716)^2}} = 3.84.$$

$t_{0.01} = 2.624$ والمستخرجة من جدول توزيع t في ملحق (٤) بدرجات حرية $v = n - 2 = 16 - 2 = 14$. وبما أن t تقع في منطقة الرفض $T > 2.624$. وبما أن t تقع في منطقة الرفض ، نرفض H_0 .

وبما أن ρ يقيس قوة الارتباط الخطى بين متغيرين X, Y في المجتمع فإن فرض العدم $H_0: \rho = 0$ يدل على عدم وجود ارتباط بين المتغيرين في المجتمع . في البند (١٠-٢-٧) استخدما

القيمة $t = \frac{b_1}{\sqrt{s^2 / SXX}}$ لاختبار فرض العدم $H_0: \beta_1 = 0$ هنا يمكن البت أن
 $t = r\sqrt{n-2} / \sqrt{1-r^2} = b_1 / \sqrt{s^2 / SXX}$ وهذا يعنى أن الاختبارين متكافئين . وعلى
ذلك إذا كان الاهتمام فقط بقياس قوة العلاقة بين متغيرين X, Y وليس الحصول على معادلة
الانحدار الخطى فإن اختبار $H_0: \rho = 0$ يكون اسهل من اختبار t لأنه يتطلب كمية قليلة من
الحسابات.

الأسلوب المستخدم لاختبار $H_0: \rho = \rho_0$ عندما $\rho_0 \neq 0$ لا يكافئ أي طريقة مستخدمة
في تحليل الانحدار . بفرض أن أزواج المشاهدات $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ تمثل عينة
عشوائية مأخوذة من مجتمع يتبع التوزيع الثنائي الطبيعي وإذا كانت n كبيرة وبافتراض صحة
فرض العدم فإن :

$$v = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right)$$

هي قيمة لتغير عشوائي V تقريباً يتبع التوزيع الطبيعي بمتوسط $\mu_V = \frac{1}{2} \ln \left(\frac{1+\rho_0}{1-\rho_0} \right)$ وتباين

$$\sigma_V^2 = \frac{1}{n-3} , \text{ حيث التباين لا يعتمد على } \rho , \text{ وعلى ذلك فإن :}$$

$$z = \frac{v - \frac{1}{2} \ln \left(\frac{1+\rho_0}{1-\rho_0} \right)}{1/\sqrt{n-3}} .$$

هي قيمة لتغير عشوائي Z تقريباً يتبع التوزيع الطبيعي القياسي. الجدول (١٦-١٠) يعطى
الفروض البديلة ومنطقة الرفض لكل فرض بديل عند مستوى معنوية α .

جدول (١٦-١٠)

الفروض البديلة	منطقة الرفض
$H_1: \rho \neq \rho_0$	$Z < -z_{\alpha/2} \text{ or } Z > z_{\alpha/2}$
$H_1: \rho > \rho_0$	$Z > z_{\alpha}$
$H_1: \rho < \rho_0$	$Z < -z_{\alpha}$

مثال (١٦-١٠) إذا كان لديك البيانات التالية :

$$n = 20, \quad \Sigma y_i = 690.30, \quad \Sigma y_i^2 = 29040.29, \\ \Sigma x_i y_i = 10818.56, \quad \Sigma x_i = 285.90, \quad \Sigma x_i^2 = 4409.55,$$

اختبر الفرض $0.5 < \rho < 0.8$ عند مستوى معنوية $\alpha = 0.05$

الحل . أي أننا نرغب في اختبار :

$$H_0 : \rho = 0.5,$$

$$H_1 : \rho > 0.5$$

$$\alpha = 0.05.$$

وحيث أن $r = .733$ فإن :

$$v = \frac{1}{2} \ln \left(\frac{1+.733}{1-.733} \right) = .935,$$

$$\mu_v = \frac{1}{2} \ln \left(\frac{1+.5}{1-.5} \right) = .549.$$

وعلى ذلك فإن :

$$z = \frac{v - \frac{1}{2} \ln[(1+\rho_0)/(1-\rho_0)]}{1/\sqrt{n-3}} \\ = (.935 - .549)\sqrt{17} = 1.59.$$

$z_{0.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣) . منطقة الرفض

$Z > 1.645$. وبما أن z تقع في منطقة القبول نقبل H_0 .

يمكن الحصول على $100\%(1-\alpha)$ فترة ثقة للمعلمة ρ من الصيغة التالية :

$$\frac{e^{2c_1} - 1}{e^{2c_1} + 1} \leq \rho \leq \frac{e^{2c_2} - 1}{e^{2c_2} + 1}$$

$$c_2 = v + \frac{z_{\alpha/2}}{\sqrt{n-3}}, \quad c_1 = v - \frac{z_{\alpha/2}}{\sqrt{n-3}} \quad \text{حيث أن}$$

البيانات في مثال (١٠-١٦) فإن :

$$r = 0.733, \quad v = 0.935, \quad n = 20,$$

$$c_1 = .935 - 1.96/\sqrt{17} = .460,$$

$$c_2 = .935 + 1.96 / \sqrt{17} = 1.410$$

بالتعويض في الصيغة التالية يمكن الحصول على 95% فترة ثقة للمعلمة ρ كالتالي :

$$\frac{e^{2(1.460)} - 1}{e^{2(1.460)} + 1} \leq \rho \leq \frac{e^{2(1.410)} - 1}{e^{2(1.410)} + 1}$$

والتي نحصل إلى :

$$0.43 \leq \rho \leq 0.89$$

Linear Multiple Regression (١٠-٥) الانحدار الخطي المتعدد

في الغالب تكون العلاقات الفعلية سواء الاقتصادية أو الاجتماعية أو السياسية معقدة بمثل فيها متغير واحد تابع وعدد من المتغيرات الأساسية المستقلة ومن الأمثلة العديدة على ذلك في مجال الاقتصاد نجد أن الكمية المستهلكة من سلعة ما تتأثر بسعر السلعة ذاتها علاوة على أسعار السلع البديلة وأيضا بالإضافة إلى ذوق المستهلك. كذلك كمية الإنتاج تتأثر بالعمل ورأسي المال والموارد الوسيطة وغيرها من عناصر العملية الإنتاجية. وفي مجال التأمين يوقف القسط التأميني على عمر المؤمن ودخله وقيمة الوثيقة وطول فترات التأمين .

Least Square Method (١٠-٥-١) طريقة المربعات الصغرى

الآن سوف نتناول مشكلة التقدير والتنبأ بقيمة متغير تابع بالاعتماد على فئة من المشاهدات المأخوذة من عدة متغيرات مستقلة X_1, X_2, \dots, X_p . كما في حالة الانحدار الخطي البسيط ، القيمة لكل متغير مستقل والمختارة بواسطة الباحث سوف تظل ثابتة . إذا تم اختيار عينة عشوائية من الحجم n من المجتمع فإن بيانات العينة سوف تكون على الشكل :

$$\{x_{1i}, x_{2i}, \dots, x_{pi}; y_i\}; i = 1, 2, \dots, n$$

مرة أخرى القيمة y_i تمثل قيمة لمتغير عشوائي Y_1 . نموذج الانحدار الخطي المتعدد النظري

سوف يكون على الشكل :

$$\mu_{Y|x_1, x_2, \dots, x_p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

حيث أن $\beta_0, \beta_1, \dots, \beta_p$ تمثل المعالم المطلوب تقديرها . نموذج الانحدار الخطي المتعدد المقدر

سوف يكون على الشكل :

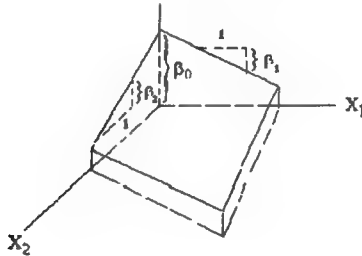
$$\hat{y}_{x_1, x_2, \dots, x_p} = b_0 + b_1 x_1 + \dots + b_p x_p$$

حيث أن b_0, b_1, \dots, b_p التقديرات المطلوب الحصول عليها للمعالم $\beta_0, \beta_1, \dots, \beta_p$. سوف تقتصر دراستنا في هذا البند على حالة وجود متغيرين مستقلين X_1, X_2 ($p=2$) . النتائج يمكن تعميمها إلى عدة متغيرات مستقلة . في حالة وجود أكثر من متغير مستقل فإن معرفتنا لنظرية المصفوفات سوف يساعدنا في عملية الحساب .

في حالة وجود متغيرين مستقلين فإن نموذج الانحدار الخطي المتعدد سوف يكون على الشكل :

$$\mu_{Y|X_1, X_2} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 ,$$

الشكل البياني للمعادلة السابقة موضح في شكل (١٠-١٤) والذي يسمى المستوى .



شكل (١٠-١٤)

ونموذج الانحدار الخطي المقدر سوف يكون على الشكل :

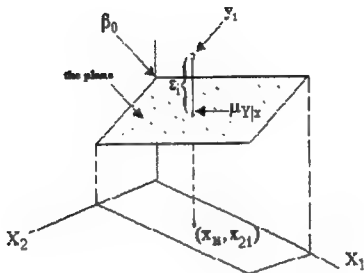
$$\hat{y}_{x_1, x_2} = b_0 + b_1 x_1 + b_2 x_2 ,$$

وكل فئة من المشاهدات تحقق العلاقة :

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + e_i, i = 1, 2, \dots, n,$$

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i, i = 1, 2, \dots, n.$$

الشكل البياني للمعادلة السابقة موضح في شكل (١٠-١٥) .



شكل (١٥-١٠)

تقديرات المربعات الصغرى b_0, b_1, b_2 يمكن الحصول عليها بحل المعادلات الخطية التالية آتيا :

$$b_0 n + b_1 \sum x_{1i} + b_2 \sum x_{2i} = \sum y_i$$

$$b_0 \sum x_{1i} + b_1 \sum x_{1i}^2 + b_2 \sum x_{1i} x_{2i} = \sum x_{1i} y_i$$

$$b_0 \sum x_{2i} + b_1 \sum x_{1i} x_{2i} + b_2 \sum x_{2i}^2 = \sum x_{2i} y_i$$

مثال (١٧-١٠) يتأثر الاستهلاك السنوي للطعام على كل من الدخل السنوي للأسرة x_1 وحجم الأسرة x_2 . أوجد معادلة الانحدار الخطي المقترنة وذلك بفرض توفر البيانات المعطاة في جدول (١٧-١٠) .

الحل . من البيانات في جدول (١٧-١٠) فإن :

$$n = 10 , \quad \sum x_{1i} = 60 , \quad \sum x_{2i} = 40 ,$$

$$\sum x_{1i}^2 = 406 , \quad \sum x_{1i} x_{2i} = 269 , \quad \sum x_{2i}^2 = 182 ,$$

$$\sum y_i = 180 , \quad \sum x_{1i} y_i = 1159 , \quad \sum x_{2i} y_i = 766 , \quad \sum y_i^2 = 3396 .$$

بالتعويض بالقيم السابقة في المعادلات الثلاثة نحصل على :

$$10 b_0 + 60 b_1 + 40 b_2 = 180$$

$$60 b_0 + 406 b_1 + 269 b_2 = 1159$$

$$40 b_0 + 269 b_1 + 182 b_2 = 766 .$$

الحل . هذه الفئة من المعادلات نحصل على :

$$b_0 = 7.918, b_1 = 2.363, b_2 = -1.024$$

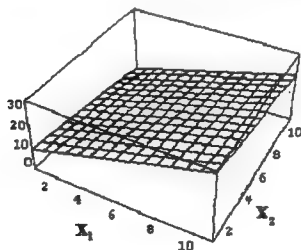
جدول (١٧-١٠)

الأمرة	الاستهلاك السنوي للطعام (بمئات الدولارات)	الدخل السنوي الصافي (بمئات الدولارات)	حجم الأسرة (عدد الأفراد في الأسرة الواحدة)
1	22	8	6
2	23	10	7
3	18	7	5
4	9	2	2
5	14	4	3
6	20	6	4
7	21	7	4
8	18	6	3
9	16	4	3
10	19	6	3

وعلى ذلك فإن نموذج الانحدار الخطي المقدر يمكن كتابته على الشكل :

$$\hat{y}_{x_1, x_2} = 7.918 + 2.363x_1 - 1.024x_2$$

والتمثيل البياني لها موضح في شكل (١٦-١٠) .



شكل (١٦-١٠)

عادة ، الخطوة الأولى بعد الحصول على معادلة الانحدار المتعدد هو اختبار ما إذا كانت هناك علاقة بين المتغير التابع وفترة المتغيرات المستقلة. يعتبر اختبار F في الانحدار المتعدد تعميم لاختبار F في حالة الانحدار الخطي البسيط في حالة متغيرين مستقلين . فرض العدم سوف يكون :

$$H_0 : \beta_1 = 0, \beta_2 = 0$$

ضد الفرض البديل : على الأقل واحد من $\beta_i : i = 1, 2$ $\beta_i \neq 0$ يعطى جدول (١٨-١٠) مجاميع المربعات ودرجات الحرية المقابلة لها .

جدول (١٨-١٠)

درجات الحرية	مجموع المربعات
n-1	$SSTO = \Sigma(y_i - \bar{y})^2$
n-3	$SSE = \Sigma(y_i - \hat{y}_i)^2$
2	$SSR = \Sigma(\hat{y}_i - \bar{y})^2$

$$\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i} \quad \text{حيث :}$$

جدول تحليل التباين معطى في جدول (١٩-١٠) .

جدول (١٩-١٠)

متوسط مجموع المربعات	مجموع المربعات	درجات الحرية	مصدر الاختلاف
$MSR = \frac{SSR}{2}$	SSR	2	الانحدار
$MSE = \frac{SSE}{n-3}$	SSE	n-3	الخطأ
	SSTO	n-1	الكلى

بافتراض صحة فرض العدم فإن :

$$f = \frac{MSR}{MSE}$$

تمثل قيمة لمطير عشوائي F يتبع توزيع F بدرجات حرية $v_1 = 2$, $v_2 = n - 3$. لمسوى

معنوية α فإن منطقة الرفض $F > f_{\alpha}(v_1, v_2)$ حيث $f_{\alpha}(v_1, v_2)$ تستخرج من الجدول

في ملحق (٦) بدرجات حرية v_1, v_2 . إذا وقعت f في منطقة الرفض نرفض H_0 .

للبينات في مثال (١٧-١٠) فرض العدم سوف يكون :

$$H_0 : \beta_1 = 0, \beta_2 = 0$$

ضد الفرض البديل : على الأقل واحد من β_i و $H_1: i=1,2$

يمكن وضع فرض العدم والفرض البديل على الشكل :

H_0 : الانحدار غير معنوي .

H_1 : الانحدار معنوي .

الحل . من البيانات في جدول (١٠-١٧) فإن جدول تحليل التباين معطى في جدول (١٠-٢٠) .

جدول (١٠-٢٠)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات
الانحدار	2	139.56725	69.78363
الخطأ	7	16.43275	2.34754
الكلي	9	156	

$$f = \frac{69.78363}{2.34754} = 29.7263.$$

$f_{0.05}(2,7) = 4.74$ والمستخرجة من جدول توزيع F في ملحق (٦) بدرجات حرية

$v_1 = 2$, $v_2 = 7$. وبما أن f تقع في منطقة الرفض نرفض H_0 .

يعتبر MSE تقدير للتباين σ^2 . وعلى ذلك فإن التقدير بنقط للمعلمة σ^2 هو $s^2 = 2.34754$

. وعلى ذلك فإن التقدير بنقطة للانحراف المعياري σ هو :

$$s = \sqrt{2.34754} = 1.5321684.$$

(١٠-٥-٣) معامل التحديد المتعدد Coefficient of Multiple Determination

معامل التحديد المتعدد ، يرمز له بالرمز R^2 هو :

$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO}.$$

في حالة وجود متغير مستقل واحد فإن R^2 يصبح معامل التحديد البسيط r^2 . يتراوح قيمة R^2

من الصفر إلى الواحد الصحيح ، أي أن :

$$0 \leq R^2 \leq 1$$

عندما $R^2 = 0$ فهذا يعنى أن $b_1 = b_2 = 0$ وعندما $R^2 = 1$ فهذا يعنى أن جميع القيم المشاهدة y_i

تقع على المستوى المقدر .

Multiple and Partial Correlation (١٠-٥-٤) الارتباط المتعدد والجزئي

الجذر التربيعي الموجب لمعامل التحديد المتعدد R^2 يسمى الارتباط المتعدد ويرمز له بالرمز

R ، حيث :

$$R = +\sqrt{R^2}$$

للمثال (١٧-١٠) فإن :

$$R = +\sqrt{0.89466} = 0.94586$$

عموماً يقاس معامل التحديد المتعدد R العلاقة بين المتغير التابع Y والمتغيرات المستقلة كلها

مجتمعة. بفرض أننا نرغب في إيجاد العلاقة بين المتغير التابع Y وأحد العوامل فقط (بفرض أن

العوامل الأخرى ثابتة) أي نحذف تأثير المتغيرات الأخرى ، وهنا نستخدم معامل الارتباط

الجزئي **coefficient of partial correlation** .

إذا كنا نرغب في إيجاد معامل الارتباط الجزئي بين المتغير Y والمتغير X_1 مع استبعاد أثر

المتغير X_2 فإن معامل الارتباط الجزئي ، يرمز له بالرمز $r_{y1.2}$ هو :

$$r_{y1.2} = \frac{r_{y1} - r_{y2} r_{12}}{\sqrt{1 - r_{y2}^2} \sqrt{1 - r_{12}^2}}$$

حيث r_{y1} هو معامل الارتباط البسيط بين المتغير Y والمتغير X_1 وبالمثل يكون r_{y2} و r_{12} .

أيضاً معامل الارتباط الجزئي بين المتغير Y والمتغير X_2 مع استبعاد أثر المتغير X_1 ، يرمز له بالرمز

$r_{y2.1}$ ، هو :

$$r_{y2.1} = \frac{r_{y2} - r_{y1} r_{12}}{\sqrt{1 - r_{y1}^2} \sqrt{1 - r_{12}^2}}$$

معامل الارتباط الجزئي يمكن أن يكون موجباً أو سالباً حيث تقع قيمته في الفترة $[-1, 1]$.

لحساب معاملات الارتباط الجزئية من مثال (١٧-١٠) نقوم أولاً بحساب معاملات الارتباط

البسيطة التالية :

$$r_{y1} = \frac{\sum x_{1i} y_i - \frac{\sum x_{1i} \sum y_i}{n}}{\sqrt{\left[\sum x_{1i}^2 - \frac{(\sum x_{1i})^2}{n} \right] \left[\sum y_i^2 - \frac{(\sum y_i)^2}{n} \right]}}$$

$$= \frac{1159 - \frac{(60)(180)}{10}}{\sqrt{\left[406 - \frac{(60)^2}{10}\right] \left[3396 - \frac{(180)^2}{10}\right]}} = 0.9325795,$$

$$r_{y2} = \frac{\sum x_{2i} y_i - \frac{\sum x_{2i} \sum y_i}{n}}{\sqrt{\left[\sum x_{2i}^2 - \frac{(\sum x_{2i})^2}{n}\right] \left[\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right]}}$$

$$= \frac{766 - \frac{(40)(180)}{10}}{\sqrt{\left[182 - \frac{(40)^2}{10}\right] \left[3396 - \frac{(180)^2}{10}\right]}} = 0.785207,$$

$$r_{12} = \frac{\sum x_{1i} x_{2i} - \frac{\sum x_{1i} \sum x_{2i}}{n}}{\sqrt{\left[\sum x_{1i}^2 - \frac{(\sum x_{1i})^2}{n}\right] \left[\sum x_{2i}^2 - \frac{(\sum x_{2i})^2}{n}\right]}}$$

$$= \frac{269 - \frac{(60)(40)}{10}}{\sqrt{\left[406 - \frac{(60)^2}{10}\right] \left[182 - \frac{(40)^2}{10}\right]}} = 0.9116072.$$

وعلى ذلك فإن :

$$r_{y1.2} = \frac{r_{y1} - r_{y2} r_{12}}{\sqrt{1 - r_{y2}^2} \sqrt{1 - r_{12}^2}}$$

$$= \frac{0.9325795 - (0.785207)(0.9116072)}{\sqrt{(1 - (0.785207)^2) \sqrt{1 - (0.9116072)^2}}} = 0.859283884.5,$$

$$r_{y2.1} = \frac{r_{y2} - r_{y1} r_{12}}{\sqrt{1 - r_{y1}^2} \sqrt{1 - r_{21}^2}}$$

$$= \frac{0.785207 - (0.9325795)(0.9116072)}{\sqrt{(1 - (0.9325795)^2) \sqrt{1 - (0.9116072)^2}}} = -0.4376576.$$

Quadratic Regression (١٠-٦) الانحدار من الدرجة الثانية

في بعض الأحيان تكون العلاقة بين متغيرين على شكل منحنى من الدرجة الثانية فعلى سبيل المثال بفرض أننا نرغب في تقدير معالم النموذج :

$$\mu_{Y|X} = \beta_0 + \beta_1 X + \beta_2 X^2$$

في الحقيقة نرغب في تقدير معالم النموذج انحدار خطي متعدد على الشكل :

$$\mu_{Y|X} = b_0 + b_1 X_1 + b_2 X_2$$

وذلك بوضع $X_1 = X$, $X_2 = X^2$ في المعادلة السابقة .

مثال (١٠-١٨) لأزواج القياسات المعطاة في جدول (١٠-٢١) أوجد الانحدار المقسرة لنموذج الانحدار من الدرجة الثانية .

جدول (١٠-٢١)

x	1	2	3	4	5	6	7	8	9	10
y	5.0	10.20	15.35	20.50	25.95	32.20	38.50	46.00	53.80	62.00

الحل . من البيانات في جدول (١٠-٢١) فإن :

$$n = 10 , \Sigma x_i = 55 , \Sigma y_i = 309.5 ,$$

$$\Sigma x_i y_i = 2218.1 , \Sigma x_i^2 = 385,$$

$$\Sigma x_i^2 y_i = 17708.2 , \Sigma x_i^3 = 3025 , \bar{y} = 30.95,$$

$$\Sigma x_i^4 = 25333 , \Sigma y_i^2 = 12831.845 ,$$

تستخدم القيم السابقة في الحصول على المعادلات التالية :

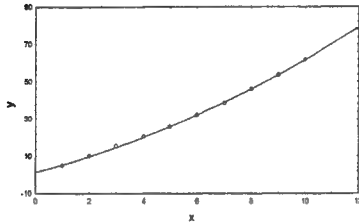
$$\begin{aligned}
 10 b_0 + 55 b_1 + 385 b_2 &= 309.5 \\
 55 b_0 + 385 b_1 + 3025 b_2 &= 2218.1 \\
 385 b_0 + 3025 b_1 + 25333 b_2 &= 17708.2
 \end{aligned}$$

بحل المعادلات السابقة آنيا يمكن إيجاد b_0, b_1, b_2 .

الحل لهذه الفئة من المعادلات هو $b_0 = 1.48083$ و $b_1 = 3.792313$ و $b_2 = 0.223674$ معادلة الانحدار المقدرة هي :

$$\hat{y}_x = 1.48083 + 3.792313x + 0.223674x^2.$$

والتمثيل البياني لها موضح في شكل (١٠-١٧).



شكل (١٠-١٧)

معامل الارتباط من الدرجة الثانية Second – Degree Correlation Coefficient

عندما تكون العلاقة بين متغيرين على شكل معادلة من الدرجة الثانية ، بمعنى أن خط الانحدار يكون على شكل منحنى (أي غير مستقيم) ، يقال في هذه الحالة أن الارتباط غير مستقيم وفي هذه الحالة لا يصلح قياس الارتباط بمعامل الارتباط الخطي البسيط r وذلك لعدم استقامة الارتباط بل يستخدم المقياس التالي :

$$\begin{aligned}
 r &= \sqrt{\frac{b_0 \sum y_i + b_1 \sum x_i y_i + b_2 \sum x_i^2 y_i - n \bar{y}^2}{\sum y_i^2 - n \bar{y}^2}} \\
 &= \sqrt{\frac{1.4808(309.5) + 3.7923(2218.1) + 0.22367(17708.2) - 10(30.95)^2}{12831.845 - 10(30.95)^2}} \\
 &= \sqrt{\frac{3251.7763}{3252.82}} = \sqrt{0.9997}
 \end{aligned}$$

$$= 0.99984.$$

أيضا معامل التحديد مازال r^2 يساوى $r^2 = 0.9997$:

تمارين :

١- أجريت تجربة لدراسة تأثير درجة الحرارة X على نتائج إحدى العمليات الكيميائية وتم الحصول على البيانات التالية (في شكل شفرة Coded) .

x	-5	-4	-3	-1	0	1	2	3	4	5
y	1	5	4	7	10	8	9	13	14	18

أوجد نموذج الانحدار الخطي المقدر .

٢- تعتبر كمية الرطوبة في منتج ما لها تأثير على كثافة المنتج النهائية. ثم مراقبة المنتج وقياس

كثافته وتسجيل البيانات التالية (في شكل شفرة Coded) .

رطوبة المنتج	x	4.7	5.0	5.2	5.2	5.9	4.7	5.9	5.2	5.9	5.6	5.0
كثافة المنتج	y	3	3	4	10	2	9	3	7	6	6	4

(أ) قدر معالم نموذج الانحدار الخطي البسيط .

(ب) أوجد 95% فترة ثقة للمعلمة β_0 .

(ج) أوجد 95% فترة ثقة للمعلمة β_1 .

(د) هل تعتقد أن المعادلة الخطية مناسبة لتوفيق هذه البيانات ؟

٣- إذا كانت تكاليف صيانة سيارات الشحن تزيد مع عمر السيارة . استخدم البيانات

التالية في (أ) تقدير معالم نموذج الانحدار الخطي البسيط (ب) هل يعتبر النموذج الخطي هو

الأفضل لتوفيق العلاقة في هذا المثال ؟

العمر بالسنوات	x	4.5	4.5	4.5	4.0	4.0	5.0	5.0	5.5	5.0	5.0	0.5	6.0	6.0
التكاليف خلال ٦ شهور	y	619	1049	1033	495	723	681	890	1522	987	1194	163	182	764

٤- في دراسة عن تأثير درجة الحرارة ، في عملية decolorizing على لون المنتج النهائي تم

الحصول على البيانات التالية :

درجة الحرارة	x	460	450	440	430	420	410	450	440	430
--------------	---	-----	-----	-----	-----	-----	-----	-----	-----	-----

0.5	0.5	0.6	0.6	0.5	0.6	0.4	0.3	0.3	y	الرد
-----	-----	-----	-----	-----	-----	-----	-----	-----	---	------

(أ) أوجد معادلة نموذج الانحدار الخطي المقدر.

(ب) أختبر معنوية كل من β_0, β_1 .

(ج) أوجد 95% فترة ثقة لكل من β_0, β_1 .

٥- في أحد أماكن بيع السيارات كانت المبيعات كالتالي :

4	1	6	5	1	1	2	3	x	عمر
28	69	15	16	68	59	10	31	y	ثمن البيع

(أ) أوجد معادلة الانحدار الخطي المقدر.

(ب) أوجد 95% فترة ثقة لكل من β_0, β_1 .

٦- إذا كان معروف أن هناك علاقة بين فترة الإصابة بمرض معين وعدد البكتريا في العضو

المصاب . تم اختيار 10 مصابا بهذا المرض وسجلت أطوال فترات إصابتهم بالمرض عند بدا

دخولهم المستشفى فتم الحصول على البيانات التالية :

6	8	4	7	6	10	7	5	10	9	x	عدد البكتريا (بالآلاف)
7	11	8	14	10	13	9	8	11	12	y	فترة الإصابة (باليوم)

أوجد معادلة الانحدار المقدر ؟

٧- فيما يلي أوزان وأطوال مجموعة من الذكور البالغين المطلوب إيجاد معادلة خط الانحدار

المقدر .

176	165	171	170	150	175	180	159	x	الطول
74	63	73	70	65	79	88	68	y	الوزن

٨- الجدول التالي يمثل الدخل والأنفاق لعبة من الأسر .

39	38	26	37	43	41	65	42	x	الدخل بمئات الجنهات
26	19	24	18	21	25	37	25	y	الإنفاق بمئات الجنهات

(أ) أوجد معادلة الانحدار المقدر.

(ت) أختبر معنوية β_0 عند مستوى معنوية 0.05.

(ج) أختبر معنوية β_1 عند مستوى معنوية 0.05.

(د) أوجد 95% فترة ثقة لكل من β_0, β_1 .

٩- الجدول التالي يوضح السن وضغط الدم لعشرة من الإناث .

السن	x	41	35	62	52	41	58	48	67	66	68
ضغط الدم	y	124	115	138	149	145	144	145	152	150	150

(أ) أوجد معادلة الانحدار المقدرة.

(ب) أوجد 99% فترة ثقة لكل من $\beta_0, \beta_1, y_{69}, \mu_{Y|170}$.

١٠- أوضحت الدراسة أن عدد العلب الصفح التي تعرض للتلف في عربة الشحن دالسة في سرعة السيارة. استخدم البيانات التالية في إيجاد :

سرعة السيارة	x	4	3	5	8	4	3	3	4	3	5	7	3	8
عدد العلب التالفة	y	63	54	86	36	65	69	28	75	53	33	168	47	52

(أ) معادلة الانحدار الخطي المقدرة.

(ب) 95% فترة ثقة لكل من β_0, β_1 .

١١- قام مركز تجاري بدراسة العلاقة بين تكاليف الإعلانات الأسبوعية و المبيعات وقد تم تسجيل البيانات والجدول التالي :

التكاليف	x	3	5	8	4	3	3	4	3	5	7	3	8
المبيعات	y	63	54	86	36	65	109	28	75	53	168	47	52

(أ) أوجد معادلة الانحدار الخطي المقدرة؟

(ب) أوجد تقدير للتباين σ^2 ؟

ج أوجد 95% فترة ثقة للمعلمة β_1 ؟

(د) أختبر فرض العلم $H_0 : \beta_1 = 0$ ضد الفرض البديل $H_1 : \beta_1 \neq 0$ وذلك عند

مستوى معنوية 0.05.

١٢- أجريت دراسة في نهر ما على عينة الثلج في الماء و تم الحصول على البيانات التالية :

كمية الثلج في أبريل	x	23.1	31.8	31.7	31.0	23.0	34.1	24.1	51.1
كمية المياه من أبريل إلى مايو	y	10.2	16.2	18.1	16.9	16.2	10.4	23.0	24.9

x كمية الطح في أبريل	36.9	30.4	25.1	12.3	36.0	31.4	21.0	27.5
y كمية المياه من أبريل إلى مايو	22.7	14.0	12.8	8.7	17.3	14.8	10.4	15.1

(أ) أوجد معادلة الانحدار الخطي المقدرة ؟

(ب) أوجد تقدير للتباين σ^2 و 95% فترة ثقة للمعلمة β_0, β_1 ؟

-١٣- البيانات التالية تمثل عدد السيارات الخاصة والتي لديها رخصة سارية خلال التسع سنوات الأخيرة في بلد ما . وقد تم تسجيل البيانات من قبل شركة لبيع السيارات وذلك للتعرف على مبيعاتها من السيارات .

x السنة	1	2	3	4	5	6	7	8	9
y عدد السيارات	8247	8919	4513	10303	10816	11228	11515	112059	12717

(أ) أوجد معادلة خط الانحدار المقدرة ؟

(ب) أوجد عدد السيارات الخاصة برخصة سارية في العام ؟

-١٤- يحتوي الجدول التالي على كمية مركب كيميائي والذي يذوب في 100 جرام من الماء عند درجات حرارة مختلفة .

$C^\circ x$	x جرام		
0	7	5	7
15	11	9	13
30	24	20	23
45	30	32	25
60	43	38	41
75	47	50	43

(أ) أوجد معادلة الانحدار الخطي المقدرة ؟

(ب) أوجد كمية المركب الكيميائي والذي يذوب في 100 جرام من الماء عند $50^\circ C$ ؟

(ج) أختبر فرض العدم H_0 : العلاقة خطية ضد الفرض البديل H_1 : العلاقة غير خطية ؟

-١٥- قام باحث بدراسة العلاقة بين الضغط (المتغير المستقل ، Kg/mm^3) ، والزمن اللازم لقطع ألواح من الصلب القاتل للصدأ (المتغير التابع ، بالساعات) في وسط ما والبيانات في الجدول التالي :

i	1	2	3	4	5	6	7	8	9	10
x_i الضغط	2.5	5	10	15	175	20	25	30	35	40
y_i الزمن القطع	63	88	55	61	62	37	38	45	46	19

(أ) أوجد شكل الانتشار ؟

(ب) أوجد معادلة خط الانحدار المقدرة ؟

(ج) أختبر معنوية كل من β_0, β_1 .

- ١٦ - من البيانات التالية أرسم شكل الانتشار وأوجد معادلة الانحدار الخطي المقدرة .

x حجم المبيعات (بالآلف دولار)	5	6	7	8	9	10
y السعر (بالآلف دولار)	74	77	82	86	92	95

وما توقعك عن السعر عندما يكون حجم المبيعات 7500 ؟

- ١٧ - البيانات التالية تم تسجيلها خلال 8 فترات .

الفترة	الوحدات المصدرة x	التكاليف الكلية y
1	10000	32000
2	20000	39000
3	30000	58000
4	40000	52000
5	50000	61000
6	60000	70000
7	7000	64000
8	90000	66000

(أ) أرسم شكل الانتشار ؟

(ب) أوجد خط الانحدار المقدر ؟

(ج) أوجد التكاليف الكلية عندما يكون عدد الوحدات المصدرة , 48750

12600 وحده.

- ١٨ - إذا كان لديك البيانات التالية :

$$n = 12, \quad \sum x_i y_i = 2000, \quad \bar{y} = 19, \quad \sum y_i^2 = 4612,$$

$$\bar{x} = 8, \quad \sum x_i^2 = 1502$$

(أ) أوجد معادلة الانحدار المقدرة ؟

(ب) أختبر الانحدار عند $\alpha = 0.05$

- ١٩ - إذا كان لديك البيانات التالية :

$$n = 25, \quad \sum (y - \bar{y})^2 = 500, \quad \sum (y_i - \hat{y}) = 125,$$

$$\sum (\hat{y}_i - \bar{y})^2 = 375$$

أوجد تقدير للتباين σ^2 ومعامل التحديد؟

٢٠- إذا كان لديك البيانات التالية :

$$n = 100, \quad \Sigma(y_i - \bar{y})^2 = 8000, \quad \Sigma(\hat{y}_i - \bar{y})^2 = 7000,$$

قدر معامل الارتباط مع العلم أن إشارة β_1 موجبة ؟

٢١- أخذت عينة عشوائية من 10 موظفين وتم الحصول على البيانات التالية :

x عدد سنوات الدراسة	12.5	14	12	16	11.0	16.0	8.0	11.2	15.0	12.7
y الأجر بمئات دولار	20.2	13.7	19.8	12.5	21.8	35.7	39.2	22.6	25.9	17.3

(أ) ارسم شكل الانتشار ؟

(ب) قدر معادلة الانحدار الخطي

(ج) أختبر فرض العدم $H_0: \beta_1 = 0$ ضد الفرض البديل $H_1: \beta_1 \neq 0$ عند

مستوى معنوية $\alpha=0.05$.

٢٢- قامت شركة بإجراء اختبار لـ 500 موظف جديد فإذا كانت x تمثل الدرجة التي

حصل عليها الموظف الجديد. وبعد ثلاثة سنوات أعطيت درجات عكسي الأداء في العمل y

الدرجة التي حصل عليها . فإذا كان لديك البيانات التالية :

$$\bar{x} = 100, \quad s_x = 10, \quad \bar{y} = 130, \quad s_y = 20, \quad r = 0.7$$

حيث S_x, S_y تمثل الانحراف المعياري من العينة لقيم x, y على التوالي :

(أ) أوجد معادلة الانحدار المقدرة ؟

(ب) أوجد قيمة y عندما $x = 90$ و $x = 125$ ؟

٢٣- قام باحث بالحصول على بيانات من 1000 مريض نفسي وذلك لمدة 5 سنوات ، فإذا

كان x يمثل الدرجة التي حصل عليها الشخص في بداية العلاج ، y الدرجة التي حصل عليها

بعد تلقى العلاج باستخدام البيانات التالية :

$$\Sigma x_i = 3000, \quad \Sigma y_i = 5000, \quad \Sigma x_i y_i = 3000$$

$$\Sigma x_i^2 = 14000.$$

(أ) أوجد معادلة الانحدار المقدرة ؟

(ب) أوجد قيمة y عندما $x=4$ ؟

(ت) إذا كانت قيمة $s_y=10$ أوجد معامل الارتباط r.

-٢٤- للبيانات التالية :

$$\bar{x} = 200, \bar{y} = 90, r = -0.9, s_x = 9, s_y = 5$$

أوجد معادلة الانحدار المقدرة وقتياً بقيمة y عندما $x = 164$.

-٢٥- في مكتب للشرطة تم إجراء دراسة للعلاقة بين عدد الجرائم في اليوم وأعلى درجة حرارة في اليوم . اختبرت عينة عشوائية في 10 أيام وتم الحصول على البيانات التالية .

x أعلى درجة حرارة	12	89	40	52	75	60	50	20	32	90
y عدد الجرائم	2	15	6	8	14	12	7	3	5	16

(أ) أوجد معادلة الانحدار المقدرة؟

(ب) أختبر فرض العدم $H_0: \beta_1 = 0$ ضد الفرض البديل $H_1: \beta_1 \neq 0$ عند مستوىمعنوية $\alpha = 0.05$.

-٢٦- الجدول التالي يعطى العمر لأحد النباتات (بالأسابيع) وطوله بالسنتيمتر .

العمر بالأسبوع x	1	2	3	4	5	6	7
الطول بالسنتيمتر y	5	13	16	33	23	38	40

أوجد معادلة خط الانحدار المقدرة ؟

-٢٧- أجريت تجربة لتقدير العلاقة بين درجة الحرارة ومعدل دقات القلب في الضفدعة المسماة

Rana Pipiens والبيانات معطاة في الجدول التالي :

الخوفان	1	2	3	4	5	6	7	8	9
x درجة الحرارة	2	4	6	8	10	12	14	16	18
y دقات القلب بالدقيقة	5	11	11	14	22	23	32	29	32

(أ) أوجد معادلة الانحدار المقدرة؟

(ب) أختبر فرض العدم $H_0: \beta_1 = 0$ ضد الفرض البديل $H_1: \beta_1 \neq 0$ عند مستوى

معنوية $\alpha=0.05$.

٢٨- أجريت تجربة لتقدير العلاقة بين العمر ودقات القلب (في الدقيقة) في الإنسان الذين أعمارهم تتراوح من واحد إلى 15 سنة . استخدم البيانات المعطاة في الجدول التالي في إيجاد :

(أ) معادلة الانحدار المقطرة؟

(ب) أختبر معنوية معادلة الانحدار؟

(ج) أوجد 95% فترة ثقة لكل من B_0 , B_1 .

الأنى	1	2	3	4	5	6	7	8	9	10
x العمر	1	2	3	4	5	6	7	8	9	10
دقات القلب y	111	108	108	102	99	92	93	88	90	90
الأنى	11	12	13	14	15					
x العمر	11	12	13	14	15					
دقات القلب y	88	84	83	83	82					

٢٩- يعطى الجدول التالي عدد سنوات الخبرة للأستاذ في جامعة ما والدخل السنوي بالدولار (بالآلاف):

x السنوات	5	10	15	20	25	30
y الدخل	59.0	69.0	78.0	88.0	97.5	107.0

(أ) أختبر العلاقة بين سنوات الخبرة والدخل السنوي ميانها؟

(ب) اقترح شكل العلاقة بين المتغيرين وقدر معالم النموذج؟

٣٠- يعطى الجدول التالي أعمار الزوج والزوجة بالسنوات لعينة من 6 أزواج:

x عمر الزوجة	35	25	51	25	53	42
y عمر الزوج	38	25	49	31	55	44

أوجد معامل الارتباط .

٣١- قام باحث بدراسة العلاقة بين ضغط الغاز (مقياس milliliters) ودرجة الحرارة (مقياس k°) وتم الحصول على البيانات التالية :

x درجة الحرارة	200	250	300	350	200	250	300	350
----------------	-----	-----	-----	-----	-----	-----	-----	-----

y الضغط	251	315	374	440	241	302	362	423
---------	-----	-----	-----	-----	-----	-----	-----	-----

ارسم شكل الانتشار وهل تعتقد أن النموذج الخطي مناسب لتوليف البيانات السابقة ؟

-٣٢- الجدول التالي يمثل درجات امتحان أعمال السنة ودرجات الامتحان النهائي لعينة عشوائية من 8 طلبة .

x درجة أعمال السنة	75	49	70	71	80	93	95	98
y درجة الامتحان النهائي	80	65	77	33	40	84	98	98

(أ) أوجد معادلة الانحدار المقترنة؟

(ب) أوجد الدرجة النهائية لطالب إذا كانت درجته في أعمال السنة 83.

-٣٣- تعمل آلة عند سرعات مختلفة ولكن السرعة العالية تؤثر على عمر يد المقطاب البيانات التالية تمثل أعمار يد المقطاب عند دورات مختلفة من الآلة.

x عدد الدورات	18	20	43	20	23	26	26	28	31	32	32	40	41	42
y عمر يد المقطاب	162	54	69	171	162	138	140	129	125	106	97	95	105	109

(أ) ارسم شكل الانتشار ؟

(ب) قدر عمر يد المقطاب عندما تكون عدد دورات الآلة 30 دورة في الدقيقة ؟

-٣٤- أعطى اختبار في المعلومات العامة يتكون من 100 سؤال إلى 20 طالب في أعمار مختلفة وكانت النتائج كالتالي :

الطالب	العمر		عدد الإجابات الصحيحة
	سنوات	شهر	
A	16	8	40
B	16	2	45
C	17	9	47
D	17	1	46
E	18	6	67
F	18	7	45
G	19	10	53
H	20	4	54

المطلوب (أ) رسم شكل الانتشار ؟

(ب) إيجاد معادلة الانحدار الخطي المقترنة ؟

(ج) أوجد باستخدام خط الانحدار المقدّر عدد الإجابات الصحيحة إذا كان عمر الطالب 17 سنة .

٣٥- أجرى ثلاثة باحثين في الجيولوجيا في القطب الشمالي بحث أستمّر خمسة فصول شتوية وذلك لدراسة العلاقة بين عدد الأيام التي تنخفض فيها درجة الحرارة عن 50°F وطول القسرون لحيوان الموطن (حيوان ضخم من الأباتل). أوجد معامل الارتباط الذي يعتمد على هذه البيانات .

x عدد الأيام التي تنخفض درجة الحرارة عن 50°F	30	20	10	30	10
متوسط الطول لقرون الحيوان (بالمتري)	0.9	0.8	0.5	1.0	0.8

٣٦- أحسب معامل الارتباط الخطي بين الفئات التالية من الأرقام

درجات اختبار ذكاء	97	121	84	105	93	126	109	97	112	116	86	103
الدخول الأسبوعي عند العسر 23	20	16	19	18	18	17	22	13	14	9	20	21

٣٧- الجدول الآتي بين طول الجمجمة X وعرضها Y بالمليمتر والمطلوب إيجاد معامل الارتباط الخطي .

x الطول	63	80	70	76	66	79	73	72	58	71
y العرض	40	42	45	38	39	46	42	37	39	35

٣٨- الجدول التالي بين المبيعات اليوم بالجنبة X لعشرة عمال في متجر ومدة خدمتهم Y بالسنين أوجد معامل الارتباط الخطي .

x	9	10	12	9	10	9	6	5	4	5
y	7	10	11	10	4	8	9	4	2	5

٣٩- الجدول التالي يعرض عدد العدسات التي تنتجها إحدى المصانع وتكلفة العدسة الواحدة بالدولار .

x عدد العدسات	1	3	5	10	12
y تكلفة العدسة	20	15	10	7	5

أوجد معامل الارتباط؟

٤٠- اختيرت عينة عشوائية من 12 ورقة من شجرة ما وتم قياس الطول والعرض لكل ورقة إلى أقرب ملليمتر . البيانات معطاة في الجدول التالي :

الورقة	1	2	3	4	5	6	7	8	9	10	11	12
العرض	35	21	25	35	26	40	35	40	35	42	23	25
الطول	55	44	46	60	55	57	64	68	5	61	46	44

(أ) أوجد معامل الارتباط ؟

(ب) أختبر معنوية معامل الارتباط المجتمع ρ ؟

أوجد معامل الارتباط؟

(ت) أختبر معنوية معامل الارتباط المجتمع ρ ؟

٤١- الجدول التالي يبين درجات مجموعة مكونة من 10 طلاب في كل من مساتي الإحصاء والرياضيات في أحد الامتحانات لأعمال الفصلين .

x الإحصاء	17	13	8	17	14	10	7	17	8	9
y الرياضيات	18	14	6	16	14	9	10	13	7	8

(أ) أوجد معامل الارتباط ؟

(ب) أوجد 95% فترة ثقة للمعلمة ρ ؟

٤٢- الجدول التالي يعطي أطوال الأب X وأطوال ابنه الأصغر عند بلوغه سن معين Y (البيانات مقاسة لأقرب بوصة) .

x	68	64	70	72	69	74
y	67	68	69	73	66	70

أوجد معامل الارتباط الخطي البسيط ؟

٤٣- لدراسة العلاقة بين عدد الكيلوجرامات التي يفقدتها شخص في برنامج لإنقاص الوزن وعدد الأسابيع التي يقضيها لإنقاص الوزن ، اختيرت عينة عشوائية من 5 أشخاص ممن يتبعون هذا البرنامج الغذائي وتم الحصول علي البيانات التالية :

x	3	2	1	4	5
y	6	5	4	9	11

أوجد معامل الارتباط الخطي ؟

٤٤- لدراسة العلاقة بين المدة اللازمة لتدريب العامل في مصنع ما وعمر العامل تم الحصول علي البيانات التالية :

x العمر (بالسنوات)	18	20	21	27	23	34	24	42	38	44
y المدة اللازمة للتدريب	8	5	6	8	7	11	8	10	6	8

أحسب معامل الارتباط الخطي البسيط ؟

٤٥- البيانات التالية تعطي متوسط درجة الحرارة اليومية وكمية الكهرباء المستهلكة وذلك خلال ثمانية أيام اختبرت في سنة ما :

x الحرارة (F°)	37	32	35	40	40	44	42	48
y استهلاك الكهرباء Megawatt Hours	3.7	3.8	3.7	3.6	3.7	3.4	3.4	3.3

أحسب معامل الارتباط الخطي البسيط ؟

٤٦- لدراسة العلاقة بين عدد الموظفين في شركة للخدمات وعدد الطلبات المتدفقة على الشركة تم الحصول على البيانات التالية :

عدد الموظفين	1	2	3	5	8	12	15
عدد الطلبات	10	15	20	25	30	35	40

(أ) أوجد نموذج الانحدار الأسى المقدر ؟

(ب) أوجد عدد الموظفين عندما تكون عدد الطلبات 45 ؟

٤٧- لدراسة العلاقة بين النسبة المئوية للمواطنين عن العمل والتغير في الأجور خلال عشر سنوات تم الحصول على البيانات التالية :

النسبة % x العاطلين عن العمل	1 11.6	2 12.2	3 12.5	4 11.7	5 11.6	6 12.1	7 12.6	8 11.7	9 11.5	10 11.6
y المتغير في الأجور	5.0	3.2	2.7	2.1	4.1	2.7	2.9	4.6	3.5	4.4

أوجد معامل الارتباط الخطي البسيط ؟

٤٨- أجريت تجربة لاختبار تأثير نوع معين من العقاقير على سريان الإدرينالين في حيوان مد. إذا كان Y يمثل عدد المملوجرامات من العقار اللازمة لإحداث سريان من الأدرينالين مقداره X وإذا كانت العلاقة التي تربط بين X, Y معطىها معادلة على الشكل $\mu_{Y|X} = b a^x$ حيث a, b ثابتان. أوجد تقديرات لمعلمتين a, b إذا كان لديك البيانات التالية :

x	1	2	3	4	5	6
y	3.1	6.2	11.3	22.0	48.0	92.0

٤٩- أجريت تجربة على أحد الأنواع الجديدة من السيارات لتحديد مسافة التوقف للسيارة عند سرعات مختلفة ، وكانت نتائج التجربة كما يلي :

السرعة x بالميل/ساعة	20	30	40	50	60	70
y مسافة التوقف بالقدم	50	85	130	200	280	290

إذا كانت العلاقة بين السرعة ومسافة التوقف هي على الشكل :

$$\mu_{Y|X} = a_0 x^{b_0}$$

استخدم البيانات السابقة في تقدير معالم النموذج ثم قدر مسافة التوقف لسيارة سرعتها 53 ميلا في الساعة .

٥٠- البيانات التالية تمثل أعداد السكان بالمليون في مدينة ما خلال الفترة من 1972-1980.

لذا كانت العلاقة بين السنوات وأعداد السكان على الشكل :

$$\mu_{Y|X} = a_0 x^{b_0}$$

أوجد معادلة الانحدار المقدرة ؟

العام x	1970	1971	1972	1973	1974	1975
y عدد السكان بالمليون	1.010	1.050	1.060	1.080	1.111	1.156
العام x	1976	1977	1978	1979	1980	
y عدد السكان بالمليون	1.155	1.170	1.201	1.230	1.330	

٥١- في دراسة أجريت على إحدى أنواع الطليبات وجد أن حجم المخ يتغير مع وزن الجسم من فرد لآخر وأن العلاقة بين حجم المخ ووزن الجسم على الشكل :

$$\mu_{Y|X} = a_0 X^{b_0}$$

استخدم البيانات التالية لإيجاد معادلة الانحدار المقدرة ؟

وزن الجسم x	30	35	37	40	41	44	46	47	49	52	54
حجم المخ y	360	379	380	390	409	408	412	419	425	435	439

٥٢- الجدول التالي يوضح متوسط طول الجمجمة مقاساً بالملليمترات ومتوسط طول الوجه y مقاساً بالملليمترات لنوع معين من القرود. فإذا كانت العلاقة بين طول الجمجمة وطول الوجه على الشكل

$$\mu_{Y|X} = a_0 X^{b_0}$$

استخدم البيانات التالية للحصول على تقديرات لكل من a_0 , b_0 .

طول الجمجمة x	75.5	99.1	107.8	113.7	115.25	120.0
طول الوجه y	30	63.5	93.7	131.0	141.2	143.22

٥٣- البيانات التالية تمثل سعر بيع السيارة وعمر السيارة لماركة خاصة بالسنوات :

العمر بالسنوات x	1	2	2	5	5	5
سعر البيع y	2350	1690	1749	1390	981	890

إذا كانت العلاقة بين عمر السيارة وسعر البيع على الشكل : $\mu_{Y|X} = \gamma \delta^x$

(أ) أوجد تقديرات لكل من γ , δ .

(ب) أوجد سعر البيع عندما يكون عمر السيارة 4 سنوات .

٥٤- يعطي الجدول التالي الضغط لغاز ما عند قيم مختلفة من الحجم .

الحجم in^3 x	50	60	70	90	100
-----------------------	----	----	----	----	-----

7.7	24.8	40.1	50.2	63.6	lb/in^3 الضغط
-----	------	------	------	------	------------------------

فإذا كان قانون الغاز معطي بالعلاقة $x^a = c \mu_Y | x$ حيث a, c ثابت أوجد :

(أ) تقديرات المربعات الصغرى للمعالم a, c

(ب) أوجد y عندما $x = 70$.

٥٥- في دراسة عن العلاقة بين عدد الحوادث في مدينة ما خلال 10 سنوات وبسبب عدد

الإشارات تم الحصول علي البيانات التالية (من 1969 إلى 1978) .

السنوات	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
x عدد الإشارات	452	425	500	517	350	480	603	611	650	718
y عدد الحوادث	42	1	6	5	1	4	7	10	15	18

استخدم النموذج الأسّي في حساب معادلة الانحدار المقبرة ؟

٥٦- البيانات التالية تعطي عدد العمال الزراعيين في الفترة من 1968 إلى 1975 .

السنوات	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975
عدد العمال	1050	1000	900	800	750	675	500	450	425	350	300

(أ) أوجد معادلة الانحدار المقبرة تحت فرض النموذج الأسّي ؟

(ب) قمر معادلة الانحدار من اللرجة التالية ؟

(ج) قمر معادلة الانحدار الخطية ؟

(د) أي المعادلات منطقية ولماذا ؟

٥٧- إذا كان لديك البيانات التالية :

x	0	1	2	3	4	5	6	7	8	9
y	9.0	7.2	3.1	4.5	4.7	2.8	5.6	7.0	8.7	10.1

(أ) المطلوب توفيق هذه البيانات باستخدام النموذج :

$$\mu_{Y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

(ب) أوجد y عند $x = 2$.

٥٨- الجدول التالي يعطي أعداد البيتزا المباعة ومتوسط التكلفة في محل لبيع البيتزا .

x عدد الوحدات المباعة	100	110	120	130	140	150	160	170	180
y متوسط التكلفة لوحدة البيع	0.90	0.88	0.85	0.72	0.65	0.70	0.73	0.79	0.74

(أ) ارسم شكل الانتشار ؟

(ب) أوجد معادلة الانحدار المقدرة تحت فرض معادلة من الدرجة الثانية ؟

-٥٩- يعطى الجدول التالي عدد القوارب المباعة في مدينة ما خلال السنوات مسن 1970 إلى 1975 .

x السنوات	1970	1971	1972	1973	1974	1975
y عدد القوارب المباعة	100	90	85	90	75	80
x السنوات	1976	1977	1975			
y عدد القوارب المباعة	86	110	115			

أستخدم معادلة من الدرجة الثانية لتوفيق هذه البيانات وأوجد عدد القوارب المباعة سنة 1981

-٦٠- أستخدم البيانات التالية في إيجاد نموذج الانحدار الخطي المتعدد المقدر ؟

x ₁ السنوات	1	2	3	4	5	6	7	8	9	10
x ₂ عدد الإشارات	452	425	500	517	350	480	603	611	850	718
y عدد الحوادث	42	1	6	5	1	4	7	10	15	18

-٦١- يتأثر محصول القراولة بكمية الأمطار وكمية السماد المستخدم أستخدم البيانات في

الجدول التالي لتوفيق معادلة انحدار خطي متعدد باستخدام كمية الإلفطار وكمية السماد

كمطهرات مستقلة ؟

x ₁ كمية الامطار	16	22	23	13	17	25	18	20	21	19	22
x ₂ السماد بالطن	510	450	500	425	450	475	515	500	490	510	525

y	1000	450	1200	700	800	1100	1050	1150	1000	950	1300
---	------	-----	------	-----	-----	------	------	------	------	-----	------

-٦٢- تبع شركة لإنتاج السحق منتجاً من خلال بعض المراكز التجارية البيانات التالية

تعطي مبيعات الشركة (بالآلاف دولار) في 10 مراكز وعدد الوحدات المباعة (بالمائة) وعدد

المراكز التجارية أوجد معادلة الانحدار الخطي المتعدد المقدرة ؟

الموقع	1	2	3	4	5	6	7	8	9	10
y المبيعات	75	38	25	98	93	54	78	85	65	88
x ₁ عدد الوحدات المباعة	15	10	7	21	14	8	14	24	9	23
x ₂ عدد المراكز التجارية	25	29	15	25	11	13	22	14	13	11

-٦٣- للبيانات التالية :

x ₁	1	4	9	11	3	8	5	10	2	7	6
x ₂	8	2	-8	-10	6	06	0	-12	4	-2	-4
y	6	8	1	0	5	3	2	-4	10	-3	5

استخدم النموذج التالي :

$$\mu_{Y|x_1, x_2} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

في إيجاد

(أ) معادلة الانحدار الخطي المتعدد المقدرة ؟

(ب) جدول تحليل التباين واستخدام $\alpha = 0.05$ في اختبار معنوية b_0, b_1, b_2 .

(ج) معامل التحديد ؟

-٦٤- البيانات التالية تعطي بيانات لثلاثة متغيرات X, Y, Z حيث X, Y متغيرات

مستقلة :

$$\sum_{i=1}^{20} x_i = 23, \quad \sum_{i=1}^{20} y_i = 40, \quad \sum_{i=1}^{20} z_i = 67, \quad \sum_{i=1}^{20} x_i^2 = 105, \quad \sum_{i=1}^{20} y_i^2 = 294,$$

$$\sum_{i=1}^{20} x_i y_i = \sum_{i=1}^{20} x_i z_i = 290, \quad \sum_{i=1}^{20} y_i z_i = 4888, \quad \sum_{i=1}^{20} z_i^2 = 815$$

استخدم البيانات السابقة في إيجاد معادلة الانحدار الخطي المقدرة ؟

٦٥- تتوقف قيمة المبيعات داخل أي جمعة استهلاكية على المربع السكني الذي توجد به الجمعية وذلك من حيث عدد السكان ومتوسط الدخل الشهري لقاطن هذا الحي بفرض أن كل جمعة يشتري منها سكان الحي الذي تتواجد به وفيما يلي بيانات لعشر جمعيات استهلاكية .

الجمعية	1	2	3	4	5	6	7	8	9	10	11	12
x_1 عدد السكان	275	175	375	200	175	260	90	300	190	50	400	270
x_2 متوسط دخل العمل	240	320	370	282	320	375	300	240	200	210	250	400
المبيعات بالآلاف جنيه	160	115	220	130	119	165	80	190	115	50	250	144

أوجد معادلة الانحدار الخطي المتعدد المقدره ؟

٦٦- في دراسة عن العلاقة بين امتصاص الماء في دقيق القمح والخواص المختلفة للدقيق وتحت فرض نموذج الانحدار خطي متعدد تم الحصول على البيانات في الجدول التالي حيث $Y(\%)$ تمثل كمية امتصاص الماء و $x_1(\%)$ كمية البروتين و $x_2(\%)$ كمية النشا الذي يتعرض للفقد التحطيم مقاس بوحدات Farrand أوجد معادلة الانحدار المقدرة ؟

x_1	8.5	8.9	10.6	10.2	9.8	10.8	11.6	12.0	12.5	10.4
x_2	2	3	3	20	22	20	31	32	31	28
y	30.9	32.7	36.7	41.9	40.9	42.9	46.3	47.2	44.0	47.7
x_1	12.2	11.9	11.3	13.0	12.9	12.0	12.9	13.1	11.4	13.2
x_2	36	28	30	27	24	25	28	28	32	28
y	43.9	46.8	46.2	47.0	46.8	45.9	48.8	46.2	47.8	49.2

الفصل الحادي عشر

تحليل التباين

Analysis of Variance

(١-١٩) مقدمة Introduction

ذكرنا في الفصل التاسع اختبار t والذي يخلص الفرق بين متوسطي مجتمعين وذلك تحت شروط معينة. في كثير من الأحيان يحتاج الباحث إلى مقارنة متوسطات ثلاثة مجتمعات فأكثر. فعلي سبيل المثال إذا كان لدينا أربع طرق للتعليم A, B, C, D يحوي الواحد منها كسل الأطفال الذين يتلقون تعليمهم بإحدى هذه الطرق والمطلوب مقارنة متوسطات المعرفة المكتسبة في كل من الطرق المختلفة. يمكن استخدام اختبار t لمقارنة متوسطي مجتمعين لكل زوج من المجتمعات الأربعة، أي استخدام اختبار t لمقارنة الطريقة A بالطريقة B ثم استخدامه مرة أخرى لمقارنة الطريقة A بالطريقة C وهكذا، إلا أن هذه الطريقة لها مشاكل كثيرة منها:

(أ) غير عملية حيث يزداد عدد المقارنات بسرعة كلما زاد عدد المجتمعات فمثلا في المثال

السابق نحتاج لإجراء اختبار t ستة مرات لأن $6 = \frac{4!}{2!2!} = \binom{4}{2}$. بصورة عامه

$$\text{عدد المقارنات الزوجية لعدد } k \text{ من المتوسطات يساوي } \binom{k}{2} = \frac{k!}{2!(k-2)!}$$

(ب) زيادة احتمال الوقوع في خطأ من النوع الأول أي رفض فرضي المسمى وهو

صحيح وذلك لأن عدد المقارنات الزوجية ومستوى المعنوية يرتبطان باحتمال

الوقوع في خطأ من النوع الأول من خلال العلاقة التالية: $1 - (1 - \alpha)^k$ حيث k

هي عدد المقارنات الزوجية و α مستوى المعنوية والذي سوف يحدد عند إجراء

مقارنة واحدة فقط. وعلى ذلك إذا كانت $k=6$ ومستوى المعنوية $\alpha=0.05$ ،

والذي يحدد لكل مقارنة زوجية، فإن احتمال الوقوع في خطأ من النوع الأول هو:

$$1 - (1 - \alpha)^k = 1 - 0.95^6 = 1 - 0.73509 = 0.26491.$$

أي ما يقرب من خمسة أمثال مستوى المعنوية $\alpha=0.05$ والذي سوف يحدد عند مقارنة

واحدة فقط للمتوسطات المستقلة في آن واحد. لحسن الحظ فإنه يمكن التغلب على المشاكل

السابقة، ومشاكل أخرى، باستخدام اختبار إحصائي يسمى تحليل التباين والذي يعتبر

واحد من أكثر الطرق الإحصائية استخداما. سوف نوضح أسلوب تحليل التباين بالمثال

التالي. إذا أجريت تجربة زراعية لدراسة تأثير الأوقات المختلفة للزراعة (فبراير - مارس -

نوفمبر - أكتوبر) على إنتاجية محصول القصب وإذا كان اهتمامنا هو اختبار فرض العدم أن

متوسط إنتاجية محصول القصب واحد للأوقات المختلفة. يعتمد أسلوب تحليل التباين، في هذه

الحالة، على تجزئة الاختلاف الكلي للمشاهدات إلى مكونين هما معنى يستخدمان في قياس

المصادر المختلفة للاختلاف . المكون الأول يقيس الاختلاف الذي يرجع إلى خطأ التجربة والثاني يقيس الاختلاف الذي يرجع إلى خطأ التجربة بالإضافة إلى الاختلاف الذي يرجع إلى أوقات الزراعة الأربعة . عندما يكون فرض العدم صحيح ، أي أن متوسط إنتاجية محصول القصب واحدة للأوقات المختلفة ، فإن كلا من المكونين سوف يمدونا بتقديرين مستقلين خطأ التجربة ، وعلى ذلك يعتمد اختبارنا على المقارنة بين المكونين باستخدام توزيع F .

بفرض أن اهتمامنا سوف يكون في مقارنة متوسط إنتاجية محصول القصب عند أوقات مختلفة للزراعة وباستخدام ثلاثة طرق للزراعة (1,2,3) . اهتمامنا في هذه الحالة سوف يكون في اختبار ما إذا كان الاختلاف في إنتاجية محصول القصب يرجع إلى الفروق في مواعيد الزراعة أو الفروق في طرق الزراعة أو ربما الفروق في كلاهما . يعتمد تحليل التباين ، في هذه الحالة ، على تجزئة الاختلاف الكلي لإنتاجية محصول القصب إلى ثلاثة مكونات ، الأول يقيس خطأ التجربة فقط والثاني يقيس خطأ التجربة بالإضافة إلى أي اختلاف يرجع إلى مواعيد الزراعة المختلفة ، والثالث يقيس خطأ التجربة بالإضافة إلى أي اختلاف يرجع إلى طرق الزراعة المختلفة . وعلى ذلك فإن مقارنة المكون الأول والثاني سوف يمدنا باختبار الفرض أن متوسط إنتاجية محصول القصب واحدة عند مواعيد الزراعة المختلفة . بنفس الشكل يمكن اختبار الفرض أن متوسط إنتاجية محصول القصب واحدة لطرق الزراعة المختلفة عن طريق مقارنة المكون الأول والثالث .

إذا صنفت المشاهدات وفقاً لصفة (خاصية) واحدة مثل الاختلاف في طرق الزراعة أو الجنس أو العمر ... الخ فسوف يكون لدينا تصنيف أحادي **one-way classification** . أما إذا صنفت المشاهدات وفقاً لصفتين مثل أصناف القمح وأنواع الأسمدة فسوف يكون لدينا تصنيف ثنائي **two-way classification** . في البند التالية سوف نتناول طرق تحليل التباين في كلا التصنيفين .

١١-١) التصنيف الأحادي One-way Classification

بفرض أن عينات عشوائية من الحجم n تم اختيارها من k من المجموعات . سوف نفترض أن المجموعات التي عددها k مستقلة وتتبع توزيعات طبيعية بموسطات $\mu_1, \mu_2, \dots, \mu_k$ وتباين مشترك σ^2 . المطلوب اختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

ضد الفرض البديل :

H_1 : واحد على الأقل من μ_i يختلف عن الباقي

نفرض أن x_{ij} ترمز للملاحظة رقم j المختارة من المجتمع رقم i وأن المشاهدات تم ترتيبها في جدول (١-١١) حيث $T_{i.}$ ترمز لمجموع كل المشاهدات في العينة المختارة من المجتمع رقم i و $\bar{x}_{i.}$ ترمز لمتوسط كل المشاهدات في العينة المختارة من المجتمع رقم i و $T_{..}$ ترمز لمجموع كل المشاهدات التي عددها nk و $\bar{x}_{..}$ ترمز لمتوسط كل المشاهدات التي عددها nk .

جدول (١-١١)

	المجتمعات				
	1	2 ...	i ...	k	
	x_{11}	$x_{21} \dots$	$x_{i1} \dots$	x_{k1}	
	x_{12}	$x_{22} \dots$	$x_{i2} \dots$	x_{k2}	
	\vdots	\vdots	\vdots	\vdots	
	x_{1n}	$x_{2n} \dots$	$x_{in} \dots$	x_{kn}	
المجموع	T_1	$T_2 \dots$	$T_i \dots$	T_k	$T_{..}$
المتوسط	\bar{x}_1	$\bar{x}_2 \dots$	$\bar{x}_i \dots$	\bar{x}_k	$\bar{x}_{..}$

يمكن التعبير عن كل مشاهدة وفقا للنموذج الرياضي التالي :

$$x_{ij} = \mu_i + \epsilon_{ij},$$

حيث ϵ_{ij} يقبض انحراف الملاحظة رقم j في العينة رقم i عن متوسط المجتمع رقم i

وبوضع $\mu_i = \mu + \alpha_i$ حيث :

$$\mu = \frac{\sum_{i=1}^k \mu_i}{k}$$

لأنه يمكن كتابة النموذج أعلاه على الشكل :

$$x_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

تحت شرط أن $\sum_{i=1}^k \alpha_i = 0$ حيث α_i تعبر عن تأثير المجتمع رقم i . وباستعمال النموذج

الأخير يصبح لفرض العدم $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$

مكافئ للفرض :

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_k = 0$$

ضد الفرض البديل :

واحد على الأقل من α_i لا يساوى صفراً : H_1

اختيارنا سوف يعتمد على مقارنة تقديرين مستقلين لتباين المجتمع σ^2 . يتم الحصول على التقديرين بتجزئه الاختلاف الكلي للملاحظات إلى مكونين . من المعروف أن التباين لكل المشاهدات مجتمعه في عينة واحدة من الحجم nk يعطى من الصيغة :

$$s^2 = \frac{\sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2}{nk - 1}$$

البسط في الصيغة السابقة يسمى مجموع المربعات الكلي **total sum of squares** والذي يقاس الاختلاف الكلي للملاحظات حيث :

$$\sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 = n \sum_{i=1}^k (\bar{x}_{i.} - \bar{x}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})^2.$$

ويمكن التعبير عن الحدود في العلاقة السابقة باستخدام الرموز كالتالي :

$$SSTO = SSC + SSE$$

حيث مجموع المربعات الكلي هو :

$$SSTO = \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_{..})^2 ,$$

ومجموع المربعات لمتوسطات الأعمدة **sum of squares for columns means** هو :

$$SSC = n \sum_{i=1}^k (\bar{x}_{i.} - \bar{x}_{..})^2 ,$$

ومجموع المربعات للخطأ **error sum of squares** هو :

$$SSE = \sum_{i=1}^k \sum_{j=1}^n (x_{ij} - \bar{x}_{i.})^2 ,$$

أيضا تجزئ درجات الحرية الكلية كما يلي :

$$nk-1 = k-1 + k(n-1).$$

عادة يشار لمجموع المربعات لمتوسطات الأعمدة من قبل كثير من المؤلفين بمجموع المربعات للمعالجات **treatment sum of squares** . وهذه التسمية ترجع إلى الحقيقة أن k من المجتمعات المختلفة غالباً ما تصنف تبعاً لمعالجات مختلفة وعلى ذلك فإن المشاهدات x_{ij} ($j=1,2,\dots,n$) تمثل n من المشاهدات المقابلة للمعالجة رقم i . الآن كلمة معالجة

تستخدم أكثر لتوضيح التصنيفات المختلفة سواء أسمدة مختلفة أو مصانع مختلفة أو مناطق مختلفة في مدينة ما أو محلين مختلفين .

التقدير الأول للمعلمة σ^2 ، يعتمد علي $k-1$ درجات حرية ، ويعطي من الصيغة :

$$MSC = \frac{SSC}{k-1}$$

عندما يكون H_0 صحيح ، فإن MSC سوف يكون تقدير غير متحيز للمعلمة σ^2 .

التقدير الثاني المستقل للمعلمة σ^2 يعتمد علي $k(n-1)$ درجات حرية ويعطي من الصيغة :

$$MSE = \frac{SSE}{k(n-1)}$$

يعتبر التقدير MSE غير متحيز بصرفه ، النظر عن صحة أو عدم صحة فرض العدم .

نعرف مما سبق أن التباين لكل مشاهدات العينة ، بدرجات حرية $nk-1$ ، هو :

$$s^2 = \frac{SSTO}{nk-1}$$

والذي يعتبر تقدير غير متحيز للمعلمة σ^2 عندما H_0 صحيح. النسبة :

$$f = \frac{MSC}{MSE}$$

هي قيمة لطيف عشوائي F يتبع توزيع F بدرجات حرية $v_1 = k-1, v_2 = k(n-1)$ عندما H_0 صحيح. مستوى معنوية α منطقة الرفض $F > f_{\alpha}(v_1, v_2)$ حيث $f_{\alpha}(v_1, v_2)$ تستخرج من جدول توزيع F في ملحق (٦) عند $\alpha = 0.05$ أو في ملحق (٧) عند $\alpha = 0.01$. إذا وقعت f في منطقة الرفض نرفض H_0 .

عمليا يتم أولا حساب SSC , SSTO , ثم نحصل علي SSE بطرح SSC من SSTO أي أن :

$$SSE = SSTO - SSC.$$

بإمكاننا حساب الصيغ السابقة والمعرفة لكل من SSTO و SSC بطريقة حسابية مبسطة (مناسبة للآلة الحاسبة) علي النحو التالي :

$$SSTO = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - CF ,$$

حيث $CF = \frac{T^2}{nk}$ يسمى معامل التصحيح correction factor . أيضا :

$$SSC = \frac{\sum_{i=1}^k T_i^2}{n} - CF.$$

عادةً الحسابات في تحليل التباين تلخص في جدول يسمى جدول تحليل التباين Analysis of Variance (عادةً يسمى ANOVA) والموضح في جدول (٢-١١)
جدول (٢-١١)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f الحسوبة
متوسطات الأعمدة	k-1	SSC	$MSC = \frac{SSC}{k-1}$	$\frac{MSC}{MSE}$
الخطأ	k(n-1)	SSE	$MSE = \frac{SSE}{k(n-1)}$	
الكلي	nk-1	SSTO		

مثال (١-١١) البيانات في جدول (٣-١١) تمثل الطول (مقاس بالسنتيمتر)
لنباتات تم زراعتها في ثلاثة أوساط مختلفة A, B, C (5 نباتات في كل وسط). أوجد
جدول تحليل التباين واختبر فرض العدم أن $\mu_1 = \mu_2 = \mu_3$ وذلك عند مستوى معنوية
 $\alpha=0.05$.

جدول (٣-١١)

الأوساط	A	10	14	18	15	12
	B	16	18	22	18	15
	C	15	12	8	10	13

الحل . المطلوب اختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

ضد الفرض البديل :

واحد علي الأقل من μ_i يختلف عن الباقى : H_1

$$\alpha=0.05 .$$

$f_{0.05}(2,12) = 3.89$ والمستخرجة من جدول توزيع F في ملحق (٦) عند درجات حرية
 $v_1 = 2, v_2 = 12$. منطقة الرفض $F > 3.89$.

$$SSTO = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - CF$$

$$= 10^2 + 14^2 + \dots + 10^2 + 13^2 - \frac{(216)^2}{15}$$

$$= 3304 - 3110.4 = 193.6,$$

$$SSC = \frac{\sum_{i=1}^k T_i^2}{n} - CF$$

$$= \frac{69^2 + 89^2 + 58^2}{5} - \frac{(216)^2}{15}$$

$$= 3209.2 - 3110.4 = 98.8.$$

تلخص النتائج في جدول تحليل التباين [جدول (١١-٤)].

جدول (١١-٤)

f المئوية	متوسط المربعات	مجموع المربعات	درجات الحرية	مصدر الاختلاف
6.25316 ^a	49.4	98.8	2	متوسطات الأعمدة
	7.9	94.8	12	الخطأ
		193.6	14	الكلي

وبما أن f (6.25316) تقع في منطقة الرفض فإننا نرفض H_0 ونعتبر أن هناك فروقاً معنوية بين متوسطات الأوساط المختلفة. النجمة * تعني أن الفرق معنوي عند $\alpha=0.05$. الآن بفرض أن العينات التي عددها k ذات أحجام n_1, n_2, \dots, n_k (عدم تساوى

حجوم العينات) حيث $N = \sum_{i=1}^k n_i$. الصيغ المستخدمة لحساب SSTO, SSC تعطى

كالتالي :

$$SSTO = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - CF,$$

$$SSC = \sum_{i=1}^k \frac{T_i^2}{n_i} - CF.$$

ويمكن الحصول علي SSE بطرح SSC من SSTO أى :

$$SSE = SSTO - SSC.$$

درجات الحرية سوف تصبح (N-1) لمجموع المربعات الكلية SSTO و (k-1) لمجموع مربعات متوسطات الأعمدة SSC و N-k = (N-1)-(k-1) لمجموع مربعات الخطأ.

مثال (١١-٢) أجريت تجربة للدراسة تأثير أربعة أنواع من الأدوية A, B, C, D على الشفاء من مرض معين. البيانات معطاة في جدول (١١-٥) والتي تمثل عدد الأيام اللازمة للشفاء . استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فرق معنوي بين المتوسطات عند مستوى معنوية $\alpha=0.05$.

جدول (١١-٥)

أنواع الأدوية			
A	B	C	D
3	7	3	10
4	8	2	12
3	4	1	8
5	10	2	5
	6	4	12
		2	10
		3	9
		1	

الحل . المطلوب اختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

ضد الفرض البديل :

$$H_1 : \mu_i \text{ واحد على الأقل من } \mu_j \text{ يختلف عن الباقي}$$

$$\alpha=0.05.$$

$f_{0.05}(3,20)=3.1$ والمستخرجة من جدول توزيع F في ملحق (٦) بدرجات حرية

$$F > 3.1 \text{ منطقة الرفض } . v_1 = 3, v_2 = 20$$

$$SSTO = \sum_{i=1}^k \sum_{j=1}^n x_{ij}^2 - CF,$$

$$= 3^2 + 4^2 + \dots + 10^2 + 9^2 - \frac{(134)^2}{24}$$

$$= 1030 - 748.17 = 281.83,$$

$$SSC = \sum_{i=1}^k \frac{T_i^2}{n_i} - CF$$

$$= \frac{15^2}{4} + \frac{35^2}{5} + \frac{18^2}{8} + \frac{66^2}{7} - \frac{(134)^2}{24}$$

$$964.04 - 748.17 = 215.87,$$

$$SSE = 281.83 - 215.87 = 65.96.$$

جدول تحليل التباين معطي في جدول (٦-١١) .

جدول (٦-١١)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسط الأعمدة	3	215.87	71.9567	21.818*
الخطأ	20	65.96	3.298	
الكلية	23	281.83		

وحيث أن f المحسوبة (21.818) تقع في منطقة الرفض فإننا نرفض H_0 . أى أن هناك فرق معنوي بين المتوسطات .

(٣-١١) اختبار تجانس عدة تباينات :

Test for the Equality of Several Variances

ذكرنا في البند (٢-١١) أن هناك الفراضات أساسية وضرورية لإجراء تحليل التباين وهم : أن التجمعات التي عددها k مستقلة وتتبع توزيعات طبيعية بمتوسطات $\mu_1, \mu_2, \dots, \mu_k$ وتباين مشترك σ^2 . هناك العديد من الطرق المختلفة لاختبار فرض العدم :

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

ضد الفرض البديل :

التباينات ليست كلها متساوية : H_1

اقترح Cochran (1991) Winer et al [القيمة التالية :

$$C = \frac{s^2 \text{ أكبر}}{\sum s_j^2}$$

والتي تمثل قيمة للإحصاء C وذلك تحت فرض أن H_0 صحيح. القيم الحرجة

$c_\alpha(v_1, v_2)$ للإحصاء C تستخرج من جدول Cochran في ملحق (٨) بدرجات

حرية $v_1 = k, v_2 = n - 1$ وذلك عند مستوى معنوية $\alpha = 0.05$ أو $\alpha = 0.01$.

منطقة الرفض $C > c_\alpha(v_1, v_2)$. إذا وقعت c في منطقة الرفض نرفض H_0 . بفرض

أن العينات التي عددها k ذات أحجام n_1, n_2, \dots, n_k (عدم تساوى أحجام العينات) وإذا كانت الأحجام متقاربة فيمكن استخدام أكبر n_i بدلاً من n في حساب درجات الحرية اللازمة لإيجاد $c_\alpha(v_1, v_2)$.

مثال (٣-١١) للبيانات في جدول (٥-١١) والخاصة بمثال (٢-١١) أختبر فرض العدم :

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$$

ضد الفرض البديل :

H_1 : التباينات ليست كلها متساوية :

وذلك عند مستوى معنوية $\alpha=0.05$.

الحل . الجدول (٧-١١) يعطي تباين العينة لكل معالجة وعدد المشاهدات في كل معالجة .

جدول (٧-١١)

المعالجة i	1	2	3	4
s_i^2	0.9167	5.0000	1.0714	5.9524
n_i	4	5	8	7

$$c = \frac{s^2 \text{ أكبر}}{\sum_{i=1}^k s_i^2} = \frac{5.9524}{12.9405}$$

$$= 0.459982.$$

وبما أن العينات التي عددها 4 ذات أحجام غير متساوية فسوف نأخذ $n = 8$ حيث 8 هي عدد المشاهدات في المعالجة رقم 3 (أكبر n_i) وعلى ذلك $v_1 = 4, v_2 = 8 - 1 = 7$ و $c_{0.05}(4, 7) = 0.5365$. منطقة الرفض $C > 0.5365$. وبما أن $c = 0.459982$ تقع في

منطقة القبول فإننا نقبل H_0 .

(٤-١١) اختبار دانكن للتعدد

Duncan's Multiple Range Test

إذا كانت قيمة f المحسوبة من جدول تحليل التباين غير معنوية فهنا يبدل على أن الفروق بين متوسطات المعالجات ليست فروق حقيقية وإنما تعزى لحد الصدفة ، وبالتساوي فإننا نقبل فرض العدم $H_0: \mu_1 = \mu_2 = \dots = \mu_k$. إذا كانت قيمة f معنوية فهنا يبدل

على أن بعض الفروق بين متوسطات المعالجات أو كلها معنوية ، ولكن هذا الاختبار لا يوضح لنا أى من هذه الفروق معنوية ، ولذلك فإن الباحث لا بد أن يجري عدة مقارنات بين هذه المتوسطات وهنا ما يسمى بالمقارنات المتعددة. هناك عدة طرق تستخدم لهذا الغرض . سوف تقتصر دراستنا في هذا البند على اختبار ^{توكسن} ^{توكسن} للمقارنات المتعددة . يُلخص اختبار ^{توكسن} في إيجاد عدة فروق معنوية ذات قيم متزايدة والتي تتوقف حجمها على مدى البعد بين المتوسطات بعد ترتيبها وتلخص خطوات تنفيذها على النحو التالي :

(أ) نرتب متوسطات المعالجات تنازلياً.

(ب) نوجد الخطأ المعياري للمتوسط $s_{\bar{x}} = \sqrt{\frac{MSE}{n}}$ حيث MSE هو متوسط

مجموع مربعات الخطأ والذي يعتبر تقدير للتباين σ^2 ، ونحصل عليه من جدول تحليل التباين . وإذا كانت أحجام العينات للمعالجات غير متساوية فإن اختبار ^{توكسن} ^{توكسن} يمكن يسمح باستبدال n في صيغة $s_{\bar{x}}$ بالوسط التوافقي للقيم n_1, n_2, \dots, n_k حيث الوسط التوافقي :

$$\tilde{n} = \frac{k}{\frac{1}{n_1} + \frac{1}{n_2} + \dots + \frac{1}{n_k}}$$

تحت شرط أن أحجام العينات تكون متقاربة من بعضها . هذا ويمكن استبدال n في صيغة $s_{\bar{x}}$ بالقيمة n^* حيث :

$$n^* = \frac{2}{\frac{1}{n(1)} + \frac{1}{n(k)}}$$

و أن :

$n(1)$ = حجم العينة المقابل لأصغر متوسط عينة .

$n(k)$ = حجم العينة المقابل لأكبر متوسط عينة .

(ج) نستخرج قيم $q_{\alpha}(p, v)$ (تسمى أقل مدى معنوي قياسي least significant

studentized range) من جدول ^{توكسن} ^{توكسن} للمدى المعنوي في ملحق (٩) حيث

$p = 2, 3, \dots, k$ و α هي مستوى المعنوية و v هي درجات حرية MSE.

(د) نحسب قيمة أقل مدى معنوي least significant range R_p وذلك بالنسبة لكل

$p = 2, 3, \dots, k$ على النحو التالي :

$$R_p = q_{\alpha}(p, v) s_{\bar{x}}, p = 2, 3, \dots, k.$$

(هـ) نقارن الفروق بين متوسطات المعالجات ونبدأ بمقارنة الفرق بين أكبر متوسط وأقل متوسط بالقيمة R_k ثم نقارن الفرق بين أكبر متوسط ولثاني أصغر متوسط بالقيمة R_{k-1} ونواصل هذه العملية وإلى أن تتم مقارنة كل الأزواج وعددها $\cdot \binom{k}{2} = k(k-1)/2$ إذا كان الفرق المحسوب بين متوسطين يساوى أو أعلى من R_p فيكون ذلك الفرق معنويًا. تلخص نتائج الاختبار بوضع خطوط مشتركة تحت المتوسطات التي لم تكن فروقها معنوية، مع الإبقاء على ترتيب المتوسطات تنازليًا: لتوضيح طريقة ^{تيسر} هانكن للمدى المتعدد فسوف نستخدم البيانات الخاصة بمشغال (١١-٢) ونتبع الخطوات التالية :

(أ) نرتب متوسطات المعالجات تنازلياً كالآتي :

$$\begin{array}{cccc} \bar{x}_4 & \bar{x}_2 & \bar{x}_1 & \bar{x}_3 \\ 9.43 & 7.00 & 3.75 & 2.25 \end{array}$$

(ب) من جدول تحليل التباين (جدول (١١-٦)) فإن $MSE = 3.298$ بدرجات حرية $v = 20$. نوجد الخطأ المعياري للمتوسط $s_{\bar{x}} = \sqrt{\frac{MSE}{n}}$ وبما أن أحجام المعالجات غير متساوية فإننا نحسب الوسط الوافقي للقيم n_1, n_2, \dots, n_k كالآتي :

$$\begin{aligned} \bar{n} &= \frac{k}{\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \frac{1}{n_4}} \\ &= \frac{4}{\frac{1}{4} + \frac{1}{5} + \frac{1}{8} + \frac{1}{7}} = \frac{4}{.7178571} = 5.5721, \end{aligned}$$

$$MSE = 3.298, s_{\bar{x}} = \sqrt{\frac{MSE}{\bar{n}}} = \sqrt{\frac{3.298}{5.5721}} = 0.7693.$$

يمكن تلخيص النتائج للحسابات السابقة في جدول (١١-٨) حيث قيم $q_{0.05}(p, 20)$ مستخرج من جدول دانكن في ملحق (٩) حيث $p = 2, 3, 4, v = 20$.

جدول (٨-١١)

p	2	3	4
$q_{.05}(p,20)$	2.95	3.58	3.96
R_p	2.27	2.75	3.05

ومقارنة قيم R_p بالفروق للمتوسطات المربعة نحصل على الاستنتاجات الآتية

* وبما أن $\bar{x}_4 - \bar{x}_3 = 7.18 > R_4 = 3.05$ فإننا نستنتج أن الفرق بين \bar{x}_4, \bar{x}_3 معنوي .

* وبما أن $\bar{x}_4 - \bar{x}_1 = 5.68 > R_3 = 2.75$ فإننا نستنتج أن الفرق بين \bar{x}_4, \bar{x}_1 معنوي .

* وبما أن $\bar{x}_4 - \bar{x}_2 = 2.43 > R_2 = 2.27$ فإننا نستنتج أن الفرق بين \bar{x}_4, \bar{x}_2 معنوي .

* وبما أن $\bar{x}_2 - \bar{x}_3 = 4.75 > R_3 = 2.75$ فإننا نستنتج أن الفرق بين \bar{x}_2, \bar{x}_3 معنوي .

* وبما أن $\bar{x}_2 - \bar{x}_1 = 3.25 > R_2 = 2.27$ فإننا نستنتج أن الفرق بين \bar{x}_2, \bar{x}_1 معنوي .

* وبما أن $\bar{x}_1 - \bar{x}_3 = 1.5 < R_2 = 2.27$ فإننا نستنتج أن الفرق بين \bar{x}_1, \bar{x}_3 غير معنوي .

عادة تلخص الاستنتاجات بوضع خطوط تحت المتوسطات التي ليست بينها فروق معنوية وذلك كما يلي:

$$\bar{x}_4 \quad \bar{x}_2 \quad \underline{\bar{x}_1 \quad \bar{x}_3}$$

يتضح من الشكل السابق أن $\mu_4 > \mu_1, \mu_4 > \mu_3$ و

$$\mu_2 > \mu_1, \mu_2 > \mu_3, \mu_4 > \mu_2$$

يمكن تلخيص النتائج السابقة على النحو الموضح في جدول (٩-١١) حيث وضعت كل الفروق الممكنة بين المتوسطات داخل الجدول وتمت مقارنتها بقيم R_p المناسبة . يتضح من جدول (٩-١١) أن الفروق على كل قطر قيمة من أعلى اليسار إلى أدنى اليمين لها نفس قيم p . على سبيل المثال الفروق 1.5, 3.25, 2.43 تقع على قطر واحد ولها $p = 2$. القيمة الخارجة لهذه الفروق هي آخر قيمة في العمود الأخير (2.27) . أيضا الفروق 5.68

4.75 تقع على لطر واحد ولها $p=3$ وتقارن بالقيمة 2.75 (القيمة الثانية في العمود الأخير) . أخيرا الفرق 7.18 يقارن عند $p=4$ بالقيمة الحرجة 3.05 وهي القيمة الأولى في العمود الأخير . النجمة * في الجدول تدل على أن الفرق معنوي وذلك عند استخدام $\alpha=0.05$.

جدول (٩-١١)

الموسطات	$\bar{x}_4 = 9.43$	$\bar{x}_2 = 7.00$	$\bar{x}_1 = 3.75$	$\bar{x}_3 = 2.25$	p	Rp
$\bar{x}_4 = 9.43$	---	2.43*	5.68*	7.18*	4	3.05
$\bar{x}_2 = 7.00$		---	3.25*	4.75*	3	2.75
$\bar{x}_1 = 3.75$			---	1.5	2	2.27
$\bar{x}_3 = 2.25$				---		

للسهولة يمكن تلخيص نتائج جدول (٩-١١) وذلك في جدول (١٠-١١) . نلاحظ أننا لم نرصد قيمة للفرق بين أي المتوسطين موضع المقارنة كما كنا نفعل من قبل بل رصدنا فقط نجمة .

جدول (١٠-١١)

	4	2	1	3
4		*	*	*
2			*	*
1				
3				

(١١-٥) التصنيف الثنائي ، مشاهدة واحدة في كل خلية

Two-Way Classification, Single Observation Per Cell

قد تصنف فئة من المشاهدات تبعاً لصفيتين معاً . على سبيل المثال عندما يرغب الباحث في مجال الزراعة في دراسة تأثير الطرق المختلفة للزراعة (ثلاثة طرق) وكذلك الأوساط المختلفة للزراعة (مارس وفراير ونوفمبر وأكتوبر) على إنتاجية محصول القصب . المشاهدات في هذه الحالة يمكن وضعها في جدول من ثلاثة صفوف وأربعة أعمدة حيث تمثل

الصفوف طرق الزراعة 1, 2, 3 وتمثل الأعمدة أوقات الزراعة (مارس وفبراير ونوفمبر وأكتوبر). يطلق على تقاطع أى صف مع أى عمود بالخلية وكل خلية تحتوي على مشاهدة واحدة. عموماً ، في حالة التصنيف الثنائي لمشاهدة واحدة يمكن وضع المشاهدات في جدول يتكون من r من الصفوف و c من الأعمدة كما هو موضح في جدول (١١-١١) حيث أن x_{ij} ترمز للمشاهدة في الصف رقم i والعمود رقم j . سوف نفترض أن x_{ij} قيم لمخرجات عشوائية مستقلة لها توزيعات طبيعية بمتوسط μ_{ij} وتباين مشترك σ^2 . في جدول (١١-١١) $T_{i.}$ و $\bar{x}_{i.}$ ترمز للمجموع والمتوسط على التوالي لكامل المشاهدات في الصف رقم i و $T_{.j}$ و $\bar{x}_{.j}$ ترمز للمجموع والمتوسط لكل المشاهدات في العمود رقم j و $T_{..}$, $\bar{x}_{..}$ ترمز للمجموع والمتوسط على التوالي لكل المشاهدات التي عددها rc . المتوسط لموسطات المجتمعات للصف رقم i ، $\mu_{i.}$ ، يعرف كالآتي :

$$\mu_{i.} = \frac{\sum_{j=1}^c \mu_{ij}}{c}.$$

جدول (١١-١١)

الصف	الأعمدة				المجموع	المتوسط
	1	2 ...	j ...	c		
1	$x_{11} \dots$	$x_{12} \dots$	$x_{1j} \dots$	x_{1c}	$T_{1.}$	$\bar{x}_{1.}$
2	x_{21}	$x_{22} \dots$	$x_{2j} \dots$	x_{2c}	$T_{2.}$	$\bar{x}_{2.}$
:	:	:	:	:	:	:
i	$x_{i1} \dots$	$x_{i2} \dots$	$x_{ij} \dots$	x_{ic}	$T_{i.}$	$\bar{x}_{i.}$
:	:	:	:	:	:	:
r	$x_{r1} \dots$	$x_{r2} \dots$	$x_{rj} \dots$	x_{rc}	$T_{r.}$	$\bar{x}_{r.}$
المجموع	$T_{.1}$	$T_{.2} \dots$	$T_{.j} \dots$	$T_{.c}$	$T_{..}$	$\bar{x}_{..}$
المتوسط	$\bar{x}_{.1}$	$\bar{x}_{.2}$	$\bar{x}_{.j}$	$\bar{x}_{.c}$		

وبنفس الشكل ، المتوسط لموسطات المجتمعات للعمود رقم j و $\mu_{.j}$ يعرف كالآتي :

$$\mu_{.j} = \frac{\sum_{i=1}^r \mu_{ij}}{r}.$$

والموسيط لموسطات المجتمعات التي عددها rc ، μ ، يعرف كالآتي :

$$\mu = \frac{\sum_{i=1}^r \sum_{j=1}^c \mu_{ij}}{rc}$$

لتقدير ما إذا كان جزء من الاختلاف بين المشاهدات يرجع إلى الاختلاف بين الصفوف ، فإننا نختبر فرض العدم :

$$H_0' : \mu_{1.} = \mu_{2.} = \dots = \mu_{r.} = \mu_{.}$$

ضد الفرض البديل :

$$H_1' : \mu_{i.} \text{ يختلف عن الباقى}$$

وبنفس الشكل لتقدير ما إذا كان جزء من الاختلاف بين المشاهدات يرجع إلى الأعمدة ، فإننا نختبر فرض العدم :

$$H_0'' : \mu_{.1} = \mu_{.2} = \dots = \mu_{.c} = \mu$$

ضد الفرض البديل :

$$H_1'' : \mu_{.j} \text{ يختلف عن الباقى}$$

يمكن كتابته كل مشاهدة على الشكل :

$$x_{ij} = \mu_{ij} + \epsilon_{ij}$$

حيث ϵ_{ij} يقيس انحراف قيمة المشاهدات x_{ij} عن متوسط المجتمع μ_{ij} . الشكل المفضل والشائع الاستخدام لهذه المعادلة (أو النموذج) يمكن الحصول عليه بوضع $\mu_{ij} = \mu + \alpha_i + \beta_j$ حيث α_i ترمز لتأثير الصف رقم i و β_j ترمز لتأثير العمود رقم j . سوف نفترض أن تأثير الصفوف والأعمدة تجميعي additive (سوف نشرح ذلك بالتفصيل في البند التالي). وعلى ذلك يمكن إعادة كتابة x_{ij} على الشكل :

$$x_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

وذلك تحت القيد التالية :

$$\sum_{i=1}^r \alpha_i = 0 , \quad \sum_{j=1}^c \beta_j = 0$$

وعلى ذلك :

$$\mu_{i.} = \frac{\sum_{j=1}^c (\mu + \alpha_i + \beta_j)}{c} = \mu + \alpha_i$$

$$\mu_{.j} = \frac{\sum_{i=1}^r (\mu + \alpha_i + \beta_j)}{r} = \mu + \beta_j,$$

الآن اختبار فرض العدم :

$$H'_0 : \mu_1 = \mu_2 = \dots = \mu_r = \mu,$$

ضد الفرض البديل : واحد على الأقل من μ_i يختلف عن الباقي :

يكافئ اختبار فرض العدم :

$$H'_0 : \alpha_1 = \alpha_2 = \dots = \alpha_r = 0$$

ضد الفرض البديل :

واحد على الأقل من α_i لا يساوي صفراً :

وبنفس الشكل اختبار فرض العدم :

$$H''_0 : \mu_1 = \mu_2 = \dots = \mu_c = \mu$$

ضد الفرض البديل :

واحد على الأقل من μ_j يختلف عن الباقي :

يكافئ اختبار فرض العدم :

$$H''_0 : \beta_1 = \beta_2 = \dots = \beta_c = 0$$

ضد الفرض البديل :

واحد على الأقل من β_j لا يساوي صفراً :

يعتمد الفرضين السابقين على المقارنة بين تقديرين مستقلين لعلمة التباين المشترك σ^2 .

هذين التقديرين نحصل عليهما بتجزئة مجموع المربعات الكلي للملاحظات إلى ثلاثة مكونات

كما يلي :

$$\sum_{i=1}^r \sum_{j=1}^c (x_{ij} - \bar{x}_{..})^2 = c \sum_{i=1}^r (\bar{x}_{i.} - \bar{x}_{..})^2 + r \sum_{j=1}^c (\bar{x}_{.j} - \bar{x}_{..})^2$$

$$+ \sum_{i=1}^c \sum_{j=1}^c (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2$$

ويمكن التعبير عن الحدود في العلاقة السابقة باستخدام الرموز كالتالي :

$$SSTO = SSR + SSC + SSE,$$

حيث مجموع المربعات الكلي هو :

$$SSTO = \sum_{i=1}^r \sum_{j=1}^c (x_{ij} - \bar{x}_{..})^2$$

ومجموع المربعات لمتوسطات الصفوف : sum of squares for rows means هو :

$$SSR = c \sum_{i=1}^r (\bar{x}_{i.} - \bar{x}_{..})^2,$$

ومجموع المربعات لمتوسطات الأعمدة هو :

$$SSC = r \sum_{j=1}^c (\bar{x}_{.j} - \bar{x}_{..})^2,$$

ومجموع مربعات الخطأ هو :

$$SSE = \sum_{i=1}^r \sum_{j=1}^c (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x}_{..})^2.$$

التقدير الأول للمعلمة σ^2 ، يعتمد على $r-1$ درجات حرية ويعطى كالآتي :

$$MSR = \frac{SSR}{r-1}.$$

عندما $\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$ فإن MSR يعتبر تقديراً غير متحيز للمعلمة σ^2 .

التقدير الثاني للمعلمة σ^2 يعتمد على $c-1$ درجات حرية ويعطى كالآتي :

$$MSC = \frac{SSC}{c-1}.$$

يعبر التقدير MSC بتقدير غير متحيز للمعلمة σ^2 إذا كان $\beta_1 = \beta_2 = \dots = \beta_c = 0$.

التقدير الثالث للمعلمة σ^2 والذي يعتمد على $(r-1)(c-1)$ درجات حرية مستقل عن

MSR و MSC ويعطى من الصيغة :

$$MSE = \frac{SSE}{(r-1)(c-1)},$$

وهو تقدير غير متحيز بصرف النظر عن صحة أو عدم صحة فرض العدم .

لاختبار فرض العدم H_0' فإننا نحسب النسبة :

$$f_1 = \frac{MSR}{MSE}$$

وهي قيمة لتغير عشوائي F_1 يتبع توزيع F بدرجات حرية $(r-1)$ و $(r-1)(c-1)$

وذلك عندما يكون فرض العدم صحيح . سوف نرفض فرض العدم ، عند مستوى معنوية

α ، عندما :

$$F_1 > f_{\alpha}(r-1, (r-1)(c-1)).$$

بنفس الشكل لاختبار فرض العدم H_0'' فإننا نحسب النسبة :

$$f_2 = \frac{MSC}{MSE}.$$

وهي قيمة لمعير عشوائي F_2 يتبع توزيع F بدرجات حرية $(r-1)(c-1)$ و $(c-1)$ وذلك عندما يكون فرض العدم صحيح. سوف نرفض فرض العدم ، عند مستوى معنوية α ، عندما :

$$F_2 > f_{\alpha}(c-1, (r-1)(c-1))$$

عملياً أولاً نحسب $SSTO$ و SSR و SSC ثم نحصل على SSE بطرح كل من SSR و SSC من SST . أي أن $SSE = SSTO - SSR - SSC$. عادة درجات الحرية المرتبطة بـ SSE تحسب بطرح درجات الحرية الخاص بكل من SSR و SSC من درجات الحرية الخاصة بـ $SSTO$. أي أن درجات الحرية الخاصة بـ SSE هي:

$$(r-1)(c-1) = (rc-1) - (r-1) - (c-1).$$

عادة الصيغ المفضلة لحساب مجموع المربعات تعطي كالآتي :

$$SSTO = \sum_{i=1}^r \sum_{j=1}^c x_{ij}^2 - CF,$$

$$SSR = \frac{\sum_{i=1}^r T_i^2}{c} - CF,$$

$$SSC = \frac{\sum_{j=1}^c T_j^2}{r} - CF,$$

حيث :

$$CF = \frac{T_{..}^2}{rc}.$$

جدول تحليل التباين في حالة التصنيف الثنائي بمشاهدة واحدة لكل خلية موضح في جدول (١١-١٢) .

جدول (١١-١٢)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط مجموع المربعات	f الحسوبة
متوسطات الصفوف	$r-1$	SSR	$MSR = \frac{SSR}{r-1}$	$f_1 = \frac{MSR}{MSE}$
متوسطات الأعمدة	$c-1$	SSC	$MSC = \frac{SSC}{c-1}$	$f_2 = \frac{MSC}{MSE}$
الخطأ	$(r-1)(c-1)$	SSE	$MSE = \frac{SSE}{(r-1)(c-1)}$	
المجموع	$rc-1$	SST		

مثال (١١-٤) تعطي البيانات في جدول (١١-١٣) الدرجات التي حصل عليها ستة من الطلبة في ثلاثة مقررات والمطلوب :

(أ) هل هناك تفاوت في مقدرة الطلبة ؟

(ب) هل هناك تفاوت في صعوبة المقررات (استخدم مستوى معنوية $\alpha = 0.05$) .

جدول (١١-١٣)

الطالب	المقرر		
	الرياضيات	اللغة الإنجليزية	اللغة الفرنسية
1	14	18	15
2	12	16	14
3	16	17	12
4	15	19	14
5	10	12	12
6	11	13	9

الحل .

$$H_0' : \alpha_1 = \alpha_2 = \dots = \alpha_6 = 0 \quad (أ)$$

$$H_0'' : \beta_1 = \beta_2 = \beta_3 = 0 \quad (ب)$$

(أ) واحد على الأقل من α_i لا يساوي صفراً : H_1'

(ب) واحد على الأقل من β_j لا يساوي صفراً : H_1''

$f_{.05}(5,10)=3.33$ والمستخرجة من جداول توزيع F في ملحق (٦) بدرجات حرية

$v_1 = 5$, $v_2 = 10$ أيضا $f_{.05}(2,10) = 4.1$. منطقة الرفض :

$$F_1 > 3.33 \quad (أ)$$

$$F_2 > 4.1 \quad (ب)$$

$$SSTO = \sum_{i=1}^6 \sum_{j=1}^3 x_{ij}^2 - CF$$

$$= 14^2 + 12^2 + \dots + 12^2 + 9^2 - \frac{(249)^2}{18}$$

$$= 3571 - 3444.5 = 126.5,$$

$$SSR = \frac{\sum_{i=1}^6 T_i^2}{c} - CF$$

$$= \frac{47^2 + 42^2 + 45^2 + 48^2 + 34^2 + 33^2}{3} - \frac{(249)^2}{18}$$

$$= 3515.67 - 3444.5 = 71.17,$$

$$SSC = \frac{\sum_{j=1}^3 T_j^2}{r} - CF$$

$$= \frac{78^2 + 95^2 + 76^2}{6} - \frac{(249)^2}{18}$$

$$= 3480.83 - 3444.5 = 36.33.$$

جدول تحليل التباين موضح في جدول (١١-١٤) .

جدول (١٤-١١)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الصفوف	5	71.17	14.234	$f_1=7.492$
متوسطات الأعمدة	2	36.33	18.165	$f_2= 9.561$
الخطأ	10	19	1.9	
الكلي	17	126.5		

بما أن $f_1 = 7.492$ تقع في منطقة الرفض نرفض H'_0 أي أن هناك تفاوت في مقدرة الطلبة. أيضا بما أن $f_2 = 9.561$ تقع في منطقة الرفض نرفض H''_0 أي أن هناك تفاوت في صعوبة المقررات .

(١١-٦) التصنيف التالي ، عدة مشاهدات لكل خلية

Two-Way Classification , Several Observations Per Cell

في البند (١١-٥) كان يفترض أن تأثير الصف والعمود تجميعي وهذا يكافئ أن $\mu_{ij} - \mu_{i'j} = \mu_{ij} - \mu_{i'j}$ أو $\mu_{ij} - \mu_{i'j} = \mu_{ij} - \mu_{i'j}$ ، أي i, i', j, j' . وهذا يعني أن الفرق بين متوسطات المجتمعات للعمودين j, j' متساوي لكل صف وأيضا الفرق بين متوسطات المجتمعات للصف i, i' متساوي لأي عمود . في كثير من التجارب لا يتحقق هذا الشرط وعلى ذلك فإن استخدام تحليل التباين الموضح في البند (١١-٥) يؤدي إلى استنتاجات خاطئة . للتوضيح وبالرجوع إلى التجربة الزراعية الخاصة بدراسة تأثير الأوقات المختلفة للزراعة (فبراير - مارس - أكتوبر - نوفمبر) وطرق الزراعة المختلفة 1, 2, 3 ، وبفرض أن نتيجة التجربة أوضحت أنه عند وقت الزراعة فبراير كان أعلى متوسط محصول القصب عند استخدام الطريقة 1 وأقل متوسط محصول القصب عند استخدام الطريقة 2 بينما عند وقت الزراعة مارس كان أعلى متوسط محصول القصب عند استخدام الطريقة 2 وأقل متوسط محصول القصب عند استخدام الطريقة 1 . هذه الخاصية تعرف بالتفاعل بين أوقات الزراعة وطرق الزراعة وهي تكشف عما إذا كان لأوقات الزراعة آثار مختلفة على طرق الزراعة . كما لا يكون للتفاعل أثر إذا اتضح أن طرق الزراعة موضع البحث متناظرة لدى الأوقات المختلفة للزراعة .

لاختيار الفروق بين الصفوف والأعمدة عندما لا يتحقق الشرط التجميعي ، أي في وجود تفاعل interaction بين الصفوف والأعمدة ، لا بد من إيجاد تقدير غير متحيز للمعلمة σ^2 . الفصل تقدير يمكن الحصول عليه إذا كررنا المشاهدات تحت نفس الظروف . للحصول على الصيغ العامة لتحليل التباين في هذه الحالة سوف نفترض الحالة التي تكسب فيها عدد المشاهدات (المكررات) replications في كل خلية تساوي n . يفرض أن المشاهدات التي سوف نحصل عليها من التجربة ، في هذه الحالة ، يمكن ترتيبها في جدول ، مثل الجدول (١١-١٥) ، والذي يتكون من r من الصفوف و c من الأعمدة . وكل خلية تحتوي على n من المشاهدات . يفرض أن y_{ijk} ترمز للملاحظة رقم k في الصف رقم i والعمود رقم j . المشاهدات التي عددها rcn موضحة في جدول (١١-١٥) .

جدول (١١-١٥)

الصفوف	الأعمدة					المجموع	المتوسط
	1	2	.	.	c		
1	x_{111}	x_{121}			x_{1c1}	$T_{1..}$	$\bar{x}_{1..}$
	x_{112}	x_{122}			x_{1c2}		
	.	.			.		
		
	.	.			.		
	x_{11n}	x_{12n}			x_{1cn}		
	x_{211}	x_{221}			x_{2c1}		
2	x_{212}	x_{222}			x_{2c2}	$T_{2..}$	$\bar{x}_{2..}$
	.	.			.		
		
	.	.			.		
	x_{21n}	x_{22n}			x_{2cn}		
	.	.			.		
		
r	x_{r11}	x_{r21}			x_{rc1}	$T_{r..}$	$\bar{x}_{r..}$
	x_{r12}	x_{r22}			x_{rc2}		
	.	.			.		
		
	.	.			.		
	x_{r1n}	x_{r2n}			x_{rcn}		
	.	.			.		
المجموع	$T_{.1}$	$T_{.2}$			$T_{.c}$	$T_{...}$	$\bar{x}_{...}$
المتوسط	$\bar{x}_{.1}$	$\bar{x}_{.2}$			$\bar{x}_{.c}$		

الملاحظات في الخلية رقم ij تمثل عينة عشوائية من الحجم n من مجتمع يفترض أنه يتبع توزيعاً طبيعياً بمتوسط μ_{ij} وتباين σ^2 . كل المجتمعات التي عددها rc يفترض أن لها تباين مشترك σ^2 . بقية الرموز المفيدة ، بعضها معطى في جدول (١١-١٥) ، يمكن توضيحها كالآتي :

$$T_{ij} = \text{مجموع الملاحظات في الخلية رقم } ij$$

$$T_{i.} = \text{مجموع الملاحظات في الصف رقم } i$$

$$T_{.j} = \text{مجموع الملاحظات في العمود رقم } j$$

$$T_{ren} = \text{مجموع كل الملاحظات التي عددها } ren$$

$\bar{x}_{ij.}$ = متوسط كل المشاهدات في الخلية رقم ij

$\bar{x}_{i..}$ = متوسط كل المشاهدات في الصف رقم i

$\bar{x}_{.j.}$ = متوسط كل المشاهدات في العمود رقم j

$\bar{x}_{...}$ = متوسط كل المشاهدات التي عددها rcn

كل مشاهدة في جدول (١١-١٥) يمكن كتابتها على الشكل :

$$x_{ijk} = \mu_{ij} + \epsilon_{ijk},$$

حيث ϵ_{ijk} يقيس انحراف قيمة المشاهدة x_{ijk} عن متوسط المجتمع . بفرض أن $(\alpha\beta)_{ij}$ يرمز

لتأثير التفاعل للصف i مع العمود رقم j و α_i يرمز لتأثير الصف رقم i و β_j يرمز لتأثير

العمود رقم j و μ تمثل المتوسط لكل المتوسطات فإن :

$$\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$$

وعلى ذلك :

$$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},$$

تحت القيود التالية :

$$\sum_{i=1}^r \alpha_i = 0, \sum_{j=1}^c \beta_j = 0, \sum_{i=1}^r (\alpha\beta)_{ij} = 0, \sum_{j=1}^c (\alpha\beta)_{ij} = 0$$

الفروض الثلاثة سوف نختبرها كالآتي :

$$H_0' : \alpha_1 = \alpha_2 = \dots = \alpha_r = 0, \quad (أ)$$

على الأقل واحد من α_i لا يساوي صفراً :

$$H_0'' : \beta_1 = \beta_2 = \dots = \beta_c = 0, \quad (ب)$$

على الأقل واحد من β_j لا يساوي صفراً :

$$H_0''' : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{rc} = 0, \quad (ج)$$

على الأقل واحد من $(\alpha\beta)_{ij}$ لا يساوي صفراً :

كل اختبار من الاختبارات السابقة يعتمد على تقديرات مستقلة للمعلمة σ^2 وذلك بتجزئة

مجموع المربعات الكلية للمشاهدات إلى أربعة مكونات كالآتي :

$$\begin{aligned}
& \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n (x_{ijk} - \bar{x}_{...}) = \\
& cn \sum_{i=1}^r (\bar{x}_{i..} - \bar{x}_{...})^2 + m \sum_{j=1}^c (\bar{x}_{.j.} - \bar{x}_{...})^2 \\
& + n \sum_{i=1}^r \sum_{j=1}^c (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...})^2 \\
& + \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij.})^2.
\end{aligned}$$

ويمكن التعبير عن مجموع المربعات في العلاقة السابقة باستخدام الرموز حيث مجموع المربعات الكلي هو :

$$SSTO = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n (x_{ijk} - \bar{x}_{...})^2,$$

ومجموع المربعات لميوسطات الصفوف هو :

$$SSR = cn \sum_{i=1}^r (\bar{x}_{i..} - \bar{x}_{...})^2,$$

ومجموع المربعات لميوسطات الأعمدة هو :

$$SSC = m \sum_{j=1}^c (\bar{x}_{.j.} - \bar{x}_{...})^2,$$

ومجموع المربعات للتفاعل بين الصفوف والأعمدة هو :

$$SS(RC) = n \sum_{i=1}^r \sum_{j=1}^c (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...})^2,$$

ومجموع المربعات للخطأ هو :

$$SSE = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij.})^2.$$

$$SSTO = SSR + SSC + SS(RC) + SSE \quad \text{أي أن}$$

أيضا تجزأ درجات الحرية إلى :

$$rcn - 1 = (r - 1) + (c - 1) + (r - 1)(c - 1) + rc(n - 1).$$

الآن يمكن الحصول على أربعة تقديرات غير متحيزة للمعلمة σ^2 عندما يكون H'_0, H_0'', H_0'''

صحيح وهم :

$$MSR = \frac{SSR}{r-1}, \quad MSC = \frac{SSC}{c-1},$$

$$MS(RC) = \frac{SS(RC)}{(r-1)(c-1)}, \quad MSE = \frac{SSE}{rc(n-1)}.$$

لاختبار الفرض H'_0 نحسب النسبة :

$$f_1 = \frac{MSR}{MSE},$$

والتي تمثل قيمة للمتغير العشوائي F_1 والذي يتبع توزيع F بدرجات حرية $(r-1), rc(n-1)$ عندما تكون H'_0 صحيح. نرفض H'_0 ، عند مستوى معنوية α ، عندما $F_1 > f_{\alpha}((r-1), rc(n-1))$. بنفس الشكل لاختبار الفرض H''_0 نحسب النسبة :

$$f_2 = \frac{MSC}{MSE}$$

والتي تمثل قيمة للمتغير العشوائي F_2 والذي يتبع توزيع F بدرجات حرية $(c-1), rc(n-1)$ عندما تكون H''_0 صحيح. نرفض H''_0 ، عند مستوى معنوية α ، عندما $F_2 > f_{\alpha}((c-1), rc(n-1))$. أخيراً لاختبار الفرض H'''_0 فإننا نحسب النسبة:

$$f_3 = \frac{MS(RC)}{MSE}$$

والتي تمثل قيمة للمتغير العشوائي F_3 والذي يتبع توزيع F بدرجات حرية $(r-1)(c-1), rc(n-1)$ عندما تكون H'''_0 صحيح. نرفض H'''_0 ، عند مستوى معنوية α ، عندما $F_3 > f_{\alpha}((r-1)(c-1), rc(n-1))$ عادة يتم الحصول على مجموع المربعات من الصيغ التالية :

$$SSTO = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n x_{ijk}^2 - CF,$$

حيث :

$$CF = \frac{T^2}{rcn} \quad (\text{معامل التصحيح})$$

$$SSR = \frac{\sum_{i=1}^r T_{i..}^2}{cn} - CF,$$

$$SSC = \frac{\sum_{j=1}^c T_{.j}^2}{m} - CF,$$

$$SS(RC) = \frac{\sum_{i=1}^r \sum_{j=1}^c T_{ij}^2}{n} - \frac{\sum_{i=1}^r T_{i..}^2}{cn} - \frac{\sum_{j=1}^c T_{.j}^2}{m} + CF$$

أما SSE فيمكن الحصول عليها من الصيغة التالية

$$SSE = SSTO - SSR - SSC - SS(RC).$$

الحسابات في مشكلة تحليل التباين ، في التصنيف الثنائية بعدة مشاهدات في كل خلية ، موضحة في جدول (١١ - ١٦).

جدول (١٦ - ١١)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الصفوف	r-1	SSR	$MSR = \frac{SSR}{r-1}$	$f_1 = \frac{MSR}{MSE}$
متوسطات الأعمدة	c-1	SSC	$MSC = \frac{SSC}{c-1}$	$f_2 = \frac{MSC}{MSE}$
التفاعل	(r-1)(c-1)	SS(RC)	$MS(RC) = \frac{SS(RC)}{(r-1)(c-1)}$	$f_3 = \frac{MS(RC)}{MSE}$
الخطأ	rc(n-1)	SSE	$MSE = \frac{SSE}{rc(n-1)}$	
الكلي	rcn-1	SSTO		

مثال (١١ - ٥) استعملت ثلاثة مستويات من مبيد ما لمقاومة ثلاثة أجناس من حشرة

Drosophila Pseudoobscura. تعطي البيانات في جدول (١١ - ١٧) معدلات

الوفيات خلال فترة من الزمن . وتعتمد التجربة على خمسة مشاهدات في كل خلية.

المطلوب : (أ) اختبار معنوية الفروق بين مستويات المبيد.

(ب) اختبار معنوية الفروق بين الأجناس المختلفة من الحشرات .

(ج) التفاعل بين مستويات المبيد و الأجناس (مستوى معنوية $\alpha=0.05$).

جدول (١٧ - ١١)

جدول (١١-١٧)

المستوي	الجنس		
	a ₁	a ₂	a ₃
1	60,55,52,38,31	58,53,50,35,30	37, 43, 57, 60, 66
2	44,37,54,57,65	63,59,54,38,38	59,51,53,62,71
3	46,51,63,66,74	63,44,46,66,71	51,80,68,71,55

فروض العدم سوف تكون كالتالي :

$$H'_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0, \quad (أ)$$

$$H''_0 : \beta_1 = \beta_2 = \beta_3 = 0, \quad (ب)$$

$$H'''_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{33} = 0, \quad (ج)$$

الفروض البديلة سوف تكون كالتالي :

$$H'_1 : \alpha_i \text{ لا يساوي صفراً}$$

$$H''_1 : \beta_j \text{ لا يساوي صفراً}$$

$$H'''_1 : (\alpha\beta)_{ij} \text{ لا يساوي صفراً}$$

$f_{.05}(2, 36) \approx 3.23$ و $\alpha=0.05$ المستخرجة من جدول توزيع F في ملحق (٦) . منطقة

الرفض $F_1 > 3.23$.

$F_2 > 3.23$ منطقة الرفض $f_{.05}(2, 36) \approx 3.23$

$F_3 > 2.61$ منطقة الرفض $f_{.05}(4, 36) \approx 2.61$

البيانات في جدول (١١-١٧) يمكن تلخيصها في جدول (١١-١٨) . الآن :

$$SSTO = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n x_{ijk}^2 - CF$$

$$= 60^2 + 55^2 + \dots + 71^2 + 55^2 - \frac{(2445)^2}{45}$$

$$= 139307 - 132845 = 6462,$$

$$SSR = \frac{\sum_{i=1}^r T_{i..}^2}{cn} - CF$$

$$= \frac{725^2 + 805^2 + 915^2}{15} - \frac{(2445)^2}{45}$$

$$= 134058.33 - 132845 = 1213.33,$$

$$SSC = \frac{\sum_{j=1}^c T_{.j}^2}{m} - CF$$

$$= \frac{793^2 + 768^2 + 884^2}{15} - \frac{(2445)^2}{45}$$

$$= 133341.93 - 132845 = 496.93,$$

$$SS(RC) = \frac{\sum_{i=1}^r \sum_{j=1}^c T_{ij}^2}{n} - \frac{\sum_{i=1}^r T_{i..}^2}{cn} - \frac{\sum_{j=1}^c T_{.j}^2}{m} + CF$$

$$= \frac{236^2 + 226^2 + \dots + 325^2}{5} - 134058.33$$

$$- 133341.93 + 132845 = 11.74,$$

$$SSE = SSTO - SSR - SSC - SS(RC)$$

$$= 6462 - 1213.33 - 496.93 - 11.74 = 4740.$$

جدول (١٨-١١)

المسوي	الجنس			
	♂ ₁	♂ ₂	♂ ₃	المجموع
1	236	226	263	725
2	257	252	296	805
3	300	290	325	915
المجموع	793	768	884	2445

جدول تحليل التباين معطي في جدول (١٩-١١) .

جدول (١٩-١١)

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الصفوف	2	1213.33	606.665	$f_1=4.608$
متوسطات الأعمدة	2	496.93	248.465	$f_2=1.887$
التفاعل	4	11.74	2.935	$f_3=0.022$
الخطأ	36	4740	131.666	
الكلي	44	6462		

من جدول تحليل التباين (جدول ١٩-١١) يمكن استنتاج :

(أ) نرفض H_0' لأن f_1 تقع في منطقة الرفض ، أي أن هناك فروق معنوية بين بمستويات المبيد.

(ب) نقبل H_0'' لأن f_2 تقع في منطقة القبول ، أي أنه لا يوجد فروق معنوية بين الأجناس.

(ج) نقبل H_0''' لأن f_3 تقع في منطقة القبول ، أي أنه لا يوجد تفاعل بين مستويات المبيد و أجناس الحشرة.

تمارين :

١- استخدم جدول تحليل التباين التالي ، عند مستوى معنوية $\alpha=0.05$ ، في اختبار فرض العدم : $H_0 : \mu_1 = \mu_2 = \mu_3$.

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	f المحسوبة
متوسطات الأعمدة	2	70		
الخطأ	11	30		
الكلي				

٢- في تجربة لدراسة المقاومة resistance لأربعة أنواع من الأسلاك اختبرت عينة عشوائية من 4 أسلاك من كل نوع وكانت $MSC = 2573.3$ و $MSE = 1394.2$. استخدم اختبار F عند مستوى معنوية $\alpha = 0.01$ لاختبار فرض العدم :

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

ضد الفرض البديل :

واحد على الأقل من μ_i يختلف عن الباقي : H_1

٣- يرغب طبيب متخصص في الأمراض النفسية في مقارنة ثلاث طرق لاختزال مستوى العدوانية في طلبة الجامعة. وقد تم إجراء اختبار HLT لقياس درجة العدوانية ، حيث تم اختيار الطلاب الذين لهم درجات عالية من العدوانية وعددهم 11 ، وقد تم اختيار 5 منهم للمعالجة بالطريقة A كما تم اختيار 3 للمعالجة B والثلاثة الباقين للمعالجة C . وقد استمرت المعالجة لمدة فصل دراسي . وفي نهاية الفصل الدراسي أعطي كل طالب اختبار HLT والنتائج في الجدول التالي :

A	74	63	68	80	79
B	54	74	71		
C	79	95	57		

(أ) أوجد جدول تحليل التباين ؟

(ب) اختبر معنوية الفروق بين المعالجات الثلاثة ؟

٤- أجريت تجربة للدراسة تأثير أربعة طرق على زمن الطيران لبعوضة الماريا (زمن الطيران خلال 24 ساعة)؟ الطرق هي IRN و IRC و IRS و C (معالجة المراقبة) . النتائج التي تم الحصول عليها هي :

$$\bar{x}_1 = 4.39(IRS) , \bar{x}_2 = 4.52(IRC),$$

$$\bar{x}_3 = 5.49(IRN) , \bar{x}_4 = 6.3(C),$$

$$\sum_{i=1}^4 \sum_{j=1}^{10} x_{ij}^2 = 1911.91.$$

استخدم جدول تحليل التباين لاختبار معنوية الفروق بين المعالجات المختلفة عند مستوى معنوية $\alpha=0.05$.

٥- أكمل جدول تحليل التباين التالي واختبر معنوية الفروق بين الأنواع عند مستوى معنوية $\alpha=0.01$.

الخصوبة	متوسط المربعات	مجموع المربعات	درجات الحرية	مصدر الاختلاف
			3	الأنواع
				(متوسطات الأعمدة)
	14.7137			الخطأ

الكلي	19	310.50076		
-------	----	-----------	--	--

٦- تعطي البيانات التالية أربعة قراءات لكل نوع من الطائرات حيث تمثل كل قراءة ، الزمن بالساعة ، الذي قطعه الطائرة من المدينة أ إلى المدينة ب :

نوع	1	5.0	5.5	5.9	5.7
	2	7.0	8.0	8.1	8.2
الطائرة	3	4.9	4.8	4.5	4.6
	4	7.1	7.4	7.6	7.8

(أ) أكتب جدول تحليل التباين ؟

(ب) اختبر معنوية الفروق بين الأنواع الأربعة ؟

(ج) في حالة إذا كانت الفروق معنوية أى من هذه الأنواع يختلف عن الآخر ؟

٧- البيانات التالية تمثل كمية Fe لأربعة أنواع من سبائك الحديد :

(1 Carbonate – 2 Silicate – 3 Magnetite , 4 Hematite)

	1	20.5	28.1	27.8	37.0	28.0		
		25.2	15.5	27.1	20.5	31.3		
أنواع	2	26.3	24	26.2	20.2	23.7		
السبائك		34.5	17.1	26.8	23.7	24.9		
	3	29.5	34	27.5	29.4	24.9		
		26.2	29.9	29.5	30.0	35.5		
	4	36.5	44.2	34.1	30.3	31.4	33.1	34.1
		32.9	36.3	25.5				

(أ) اختبر الفرضية $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ عند مستوى معنوية $\alpha=0.05$.

٨- لتقدير فيما إذا كان هناك فروق معنوية في كمية التروجين في ثلاثة مواقع مختلفة من

بحيرة ، أخذت 8 عينات من كل موقع . البيانات في الجدول التالي مقاسة / milligrams

100 grams

الموقع		
A	B	C
222	326	263
300	275	360
262	218	221
264	207	198
200	272	211
211	268	266
267	308	312
326	229	299

المطلوب تحليل البيانات للفروق المعنوية وإذا كانت قيمة f معنوية اختبر الفروق المعنوية بين كل أزواج متوسطات المواقع.

٩- يعيش طائر معين في ثلاثة مناطق جغرافية ، وقد اختيرت عينة عشوائية من كل منطقة في المناطق الثلاثة وتم قياس طول النقار بالمليمتر لاقرب رقم عشري والبيانات في الجدول التالي :

الموقع		
A	B	C
4.2	3.8	3.0
3.3	4.2	3.4
2.8	5.0	4.4
4.2	4.5	4.5
3.7	5.2	
4.4		
3.5		

المطلوب تحليل البيانات للمعنوية الإحصائية بين المناطق المختلفة وإذا كانت f المحسوبة معنوية اختبر الفروق المعنوية الإحصائية بين كل أزواج متوسطات المواقع .

١٠- يرغب باحث في العلوم البيولوجية في دراسة تأثير المستويات المختلفة من الإيثانول على زمن النوم. اختيرت عينة عشوائية من 5 فأر (متساوية في الوزن والعمر) لكل معالجة . وقد تم حقن كل فأر . وقد تم تسجيل سرعة حركة العينة في زمن النوم

rapid eye movement sleep time خلال فترة 24 ساعة والبيانات كما يلي :

0 g/kg	88.6	73.2	91.4	68.0	75.2
1 g/kg	63.0	53.9	69.2	50.1	71.5
2 g/kg	44.9	59.5	40.2	56.3	38.7
4 g/kg	31.0	39.6	45.3	25.2	22.7

(ب) أوجد جدول تحليل التباين؟

(ت) أختبر معنوية الفروق بين المعالجات الثلاثة ؟

(ج) أستخدم اختبار Cochran للتأكد من تجانس التباين عند مستوى معنوية

$$\alpha=0.01$$

١١- أجريت مقارنة لخمس طرق لتدريس مقرر الإحصاء الوصفي. الطرق الخمسة هي R تنفيذ برامج على الورق و R | L تنفيذ برامج على الورق + محاضرات C, تطبيق على الحاسب الآلي و C | L التطبيق مع المحاضرات و D | L تدريب ومناقشات. وقد تم اختيار عينة عشوائية من 9 طلاب لكل طريقة وبعد نهاية المقرر أعطي امتحان لمدة ساعة لكل الطلبة والنتائج في الجدول التالي :

	\bar{x}_i	s_i
L D	29.3	4.99
R	28.0	5.33
L R	30.2	3.33
C	32.4	2.94
L C	34.2	2.74

هل تدل هذه البيانات على أن هناك فروق معنوية بين متوسطات الطرق الخمسة في التدريس؟
(مستوى المعنوية $\alpha=0.05$).

١٢- لمقارنة أسعار أحجام مختلفة من الثلاجات في مدينة ما تم الحصول على النتائج التالية :

النوع	حجم العينة	\bar{x}_i	s_i^2
Frigidaire	18	123.21	18.71
GE	28	418.13	21.15
Whirlpool	19	421.72	17.8

أختبر معنوية الفروق بين متوسطات الأنواع المختلفة من الثلاجات عند مستوى معنوية $\alpha=0.05$.

١٣- الجدول التالي يعطي عدد الأميال في الجالون لثلاثة أنواع من البيرين استخدمت لمدة

5 أيام . أختبر فرض العدم أن متوسطات البيرين متساوية عند مستوى معنوية $\alpha=0.05$.

البيرين	1	11	13	14	15	12
	2	16	15	14	15	15
	3	18	16	15	16	17

١٤- يفرض أن خمسة أنواع من الخلطات الغذائية *diets* ، والمحتوية على مصادر مختلفة من الكربوهيدرات ، غذيت بما حصة مجموعات من الفئران. البيانات التالية تمثل كمية DNA في كبد كل فأر (مقاسة بالمليجرام لكل جرام من وزن الكبد).

	\bar{x}_i
نشا	2.58
سكروز	2.63
فركتوز	2.13
جلوكوز	2.03
مالتوز	2.49

حيث $\sum_{i=1}^5 \sum_{j=1}^4 x_{ij}^2 = 183.4$. هل تدل النتائج السابقة على أن هناك فروق معنوية بين المتوسطات (عند مستوى معنوية $\alpha=0.05$).

١٥- في تجربة لمقارنة ثلاثة طرق لتدريس مقرر في الرياضيات ، الطريقة الأولى كان التدريس فيها نظري - الطريقة الثانية كان التدريس فيها باستخدام شرائط الفيديو - الطريقة الثالثة كان التدريس فيها بالتطبيق على الحاسب الآلي. وقد تم اختيار عينة عشوائية من 5 طلاب لكل طريقة وطبقت كل طريقة لمدة فصل دراسي. وفي نهاية الفصل الدراسي أعطي كل طالب نفس الامتحان. الدرجات في الجدول التالي :

	86	77	94	82	86
	96	87	88	79	90
الطريقة	63	74	65	70	78
	3			2	1

أخبر معنوية الفروق بين متوسطات الطرق الثلاثة عند مستوى معنوية $\alpha=0.05$.

١٦- البيانات في الجدول التالي تمثل الدرجات النهائية في مادة الإحصاء والتي حصل عليها ثلاثة مجموعات الكلية من نفس الفرقة ثم تدريسهم على يد ثلاثة محاضرين A, B, C والمطلوب التحقق من معنوية الفروق بين هذه المجموعات عند مستوى معنوية $\alpha=0.05$.

	الجامع	
A	B	C
72	88	97
88	78	68
82	49	91
42	40	70
81	50	70
72	86	41
65	74	59

١٧- الجدول التالي يعطي الإنتاج اليومي لثلاثة آلات ، وقد تم تسجيل الإنتاج اليومي لكل آلة في فترة أربعة أيام اختبرت عشوائيا . المطلوب اختبار معنوية الفروق بين متوسطات الآلات عند مستوى معنوية $\alpha=0.05$.

الآلة		
A	B	C
110	115	120
108	114	121
107	116	122
106	117	123

١٨- بفرض أن أربعة أنواع من الفيتامينات A, B, C, D وتغذي عليها أربعة مجموعات من الأطفال متشابهين تماما (أربعة عينات عشوائية) وكانت الزيادة في وزن كل مجموعة كالآتي :

الفيتامينات			
A	B	C	D
2	2	1	4
3	1	5	3
3	2	4	1
	3	4	3

(أ) قدر متوسطات الزيادة في الوزن عند الفيتامينات الأربعة ؟

(ب) أوجد التباين لكل مجموعة ، وأجري اختبار Cochran .

(ج) استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فروق معنوية بين المتوسطات عند مستوى معنوية $\alpha=0.01$.

١٩- صممت أربعة مقاعد خاصة بالسيارات وبها حزام الأمان وذلك لإعطاء أحسن حماية من حوادث الرأس عند السرعة 35 mph أو أقل وقد أجري اختبار لحاكة الحوادث وتم الحصول على البرجات التالية :

المقاعد			
A	B	C	D
37	49	32	40
41	37	33	48
44	40	41	40
48	38	37	41
50	51	48	37
44	42	37	40

أوجد جدول تحليل التباين واختبر معنوية الفروق بين متوسطات المقاعد عند مستوى معنوية $\alpha=0.05$.

٢٠- تقدير إمكانية تأثير العامل الوراثي على ضغط الدم تم تحديد أربعة أنواع من الفئران A, B, C, D وقد اختبرت عينة عشوائية من كل نوع وتم قياس ضغط الدم لكل فأر وبيانات التجربة في الجدول التالي :

الأنواع			
A	B	C	D
83	85	88	89
81	84	94	85
85	84	91	88
59	91	92	92
85	88	95	84
92	89	88	85
91	91	87	40
80	89	91	93
79	86	90	90
82	87	93	89

المطلوب تحليل البيانات للفروق الإحصائية بين المتوسطات. وإذا كانت قيمة f معنوية أختبر للفروق الإحصائية بين كل أزواج المتوسطات (عند مستوى معنوية $\alpha=0.01$).

٢١- البيانات التالية تمثل النسبة المئوية (في المتوسط) لكحل methyl alcohol (لكل زوجة) والتي تم تحليلها من قبل أربعة معامل .

	1	85.06	82.25	84.87
	2	48.99	89.28	84.88
المعمل	3	89.48	84.72	85.1
	4	84.1	84.55	84.05

أ- أكتب جدول تحليل التباين؟

ب- هل توجد فروق معنوية بين المعامل (عند مستوى معنوية $\alpha=0.01$).

٢٢- جربت خمسة أنواع من الأسمدة على عدد من القطع المزروعة قمح . وقد طبقت المعالجة (١) على أربعة قطع والمعالجة (2) و (3) على 6 قطع والمعالجة (4) على 8 قطع. أما المعالجة (5) فقد طبقت على 3 قطع والمحصول لكل ياردة مربعة في الجدول التالي :

الأنواع				
(1)	(2)	(3)	(4)	(5)
78.9	63.5	79.1	87.0	75.9
72.3	74.1	90.3	91.0	77.2
81.1	75.5	85.6	91.2	81.2
85.7	80.8	81.4	75.3	
	71.3	74.5	79.4	
	79.4	95.3	80.7	
			82.8	
			89.6	

هل يوجد فروق بين تأثير الأنواع المختلفة من الأسمدة ؟ استعمل مستوى معنوية $\alpha=0.05$.

٢٣- للمقارنة بين أربعة أنواع من مشروب بارد (مصنعة تبعاً لمكسب اللون المضاف (بدون لون - أحمر - برتقالي - أخضر) وقد تم توزيع كل نوع عشوائياً على خمسة مواقع وسجل عدد حالات البيع لكل 1000 شخص في الموقع خلال فترة الدراسة والبيانات في الجدول التالي :

أنواع المشروب			
أخضر	برتقالي	أحمر	بدون لون
30.8	27.9	31.2	26.5
29.6	25.1	28.3	28.2
32.4	28.5	30.8	25.1
31.7	24.2	27.9	29.1
32.8	26.5	29.6	27.2

(أ) قدر متوسطات المبيعات عند الألوان المختلفة ؟

(ب) أوجد جدول تحليل التباين وأجري اختبار Cochran للتجانس ؟

(ج) استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فروق معنوية بين المتوسطات عند مستوى معنوية $\alpha=0.01$.

٢٤- قام مسئول الإنتاج في مصنع للبطاريات باختيار عمر البطارية بالساعة لأربعة أنواع من البطاريات التي تستخدم في تشغيل أجهزة الراديو المحمول . وقد تم استخدام 6 أجهزة

لكل نوع من البطاريات وتشغيل الأجهزة على موجة عالية لفترة معينة . وتم الحصول على البيانات التالية :

أنواع البطاريات			
(1)	(2)	(3)	(4)
5/5	4.7	6.1	5.5
5/0	3.9	5.7	5.1
5/2	4.3	5.0	4.3
5.3	4.5	5.3	4.1
4.8	4.3	6.3	5.1
5.0	4.0	5.8	4.2

هل هناك تفاوت في متوسط أعمار البطاريات للأنواع ؟

٢٥- للمقارنة بين أربعة أنواع من الحبوب من حيث كمية الثيامين (mg/g) قام باحث باختيار 6 عينات من كل نوع وتم الحصول على البيانات التالية :

Wheat	5.2	4.2	6.0	6.1	6.7	5.8
Barley	6.5	8.0	6.1	7.5	5.9	5.6
Maize	5.8	4.7	6.4	4.9	6.0	5.2
Watt	8.3	6.1	7.8	7.0	5.5	7.2

فهل تدل هذه البيانات على الاختلاف في متوسط الثيامين بين الأنواع الأربعة من الحبوب ؟

٢٦- أجريت تجربة لدراسة التأثير السام لثلاثة أنواع من الكيماويات A, B, C على جلد الفئران وذلك عن طريق معالجة بوصة مربعة من الجلد بالمادة الكيماوية وإعطاء درجات من 5 إلى 15 على حسب درجة التأثير على الجلد وقد تم اختبار عينة من 7 أنسجة لكل معالجة و البيانات في الجدول التالي :

A	9	6	5	7	5	6	6
B	9	9	8	8	7	7	7
C	6	5	6	8	5	5	7

(أ) اكتب جدول تحليل التباين ؟

(ب) أخير معنوية الفروق لميوسطات المعالجات ؟

٢٧- قام باحث في مجال علوم الأغذية بدراسة تأثير الكميات المختلفة من اللبن المضاف إلى عجينه الكيك (منخفض - متوسط - عالي) على حجم الكيسك المخبوزة (مقياس

milliliters per 100 grams) و البيانات في الجدول التالي :

1	351	369	381	380	370	358
2 الكيك	390	394	406	407	415	375
3	398	409	415	399		

(أ) أوجد التباين لكل مجموعة وأجرى اختبار Cochran للتجانس ؟

(ب) استخدم طريقة تحليل التباين لاختبار ما إذا كان هناك فروق معنوية بين المتوسطات عند مستوى معنوية $\alpha=0.01$.

-٢٨- البيانات التالية تم الحصول عليها من تجربة لمقارنة ثلاثة أنواع من المبيدات استخدمت في رضى أماكن مصابة بمشرة الخنفساء . كل مشاهدة تمثل عدد الوفيات من الخنفساء في منطقة محددة تحتوي على هذا المبيد .

1	11, 9, 13, 11
2	6, 28, 31, 27, 30, 33
3	19, 23, 19, 21, 20

حلل البيانات للفروق المعنوية في معدل الوفاة بين الأنواع المختلفة من المبيد وذلك عند مستوى معنوية $\alpha=0.05$.

-٢٩- لاختبار لفاعلية خمسة أنواع من الأسمدة على إنتاج النرة الصفراء وتم الحصول على البيانات التالية (نتائج المحصول مقاس بالكيلوجرامات لكل قطعه) .

المسار				
(1)	(2)	(3)	(4)	(5)
40	38	44	41	34
45	40	42	43	35
46	38	40	40	34
49	44	34	40	23

المطلوب التحقق من مواقع الفروق بين متوسطات المجموعات المختلفة عند مستوى معنوية $\alpha=0.05$.

-٣٠- لدراسة تأثير المستوى الاجتماعي لطلاب إحدى الجامعات على مستواهم العلمي ، قام باحث بتحديد ثلاثة مستويات اجتماعية واختار من كل مستوى عينة عشوائية من الطلاب وسجل المعدل التراكمي GRA لكل منهم . البيانات في الجدول التالي :

المستوي		
المستوى الأول	المستوى الثاني	المستوى الثالث
2.15	3.22	2.22
2.71	2.55	3.44
1.71	3.97	3.99
2.16	3.88	3.52
3.13	3.87	3.66

(أ) اختبر معنوية الفروق بين المستويات الاجتماعية الثلاثة وذلك عند مستوى معنوية $\alpha=0.05$.

(ب) في حالة ما إذا كانت الفروق معنوية ، أي هذه المستويات يختلف عن الآخر .

٣١- مقارنة أربعة أنواع مختلفة من الصناديق بالنسبة لقوة الضغط **Compression strength** (مقاس بالرطل) تم الحصول على البيانات التالية :

	1	655.5	788.3	734.3	721.4	679.1	699.4
	2	782.2	772.5	786.9	686.1	732.1	774.8
الصندوق	3	737.1	639.0	696.3	671.7	717.2	721.7
	4	535.1	628.7	542.4	559	586.9	52.0

اختبر معنوية الفروق بين الصناديق الأربعة عند مستوى معنوية $\alpha=0.01$.

٣٢- الجدول التالي يمثل المبيعات لثلاثة أنواع من الفطائر (المبيعات بالدولار) تم عرضها في 16 مركزاً للبيع خلال أسبوعين

	1	2161	1769	2548	1782	
الفطائر	2	2379	1419	1119	1208	1962
		1689				
	3	1479	1024	1598	4613	1913
		2215				

(أ) المطلوب تحليل هذه البيانات بأسلوب تحليل التباين عند مستوى معنوية $\alpha=0.01$.

٣٣- مقارنة أسعار ورغيف الخبز (من نوع ما) في أربعة مواقع في مدينة ما ، اختبرت عينة عشوائية من أربعة مخازن في كل منطقة وتم تسجيل المبيعات في كل مخبز لفترة معينة والبيانات في الجدول التالي :

	1	55	63	65	61
الموقع	2	58	61	55	58
	3	54	51	58	58
	4	69	70	71	77

(أ) اكتب جدول تحليل التباين ؟

(ب) اختبر معنوية الفروق بين متوسطات المواقع ؟

(ج) في حالة ما إذا كانت الفروق معنوية أي من هذه المواقع يختلف عن الآخر ؟

٣٤- أخذت عينات من الماء من مواقع مختلفة من نهر لتقدير فيما إذا كانت متساوية في كمية الأوكسجين المذاب والذي يعتبر مقياس لتلوث المياه ، وقد أخذت عينة عشوائية من كل موقع والبيانات في الجدول التالي :

	1	5.9	6.1	6.3	6.1	6.0
	2	6.3	6.6	6.4	6.4	6.5
الموقع	3	4.8	4.3	5.0	4.7	5.1
	4	6.0	6.2	6.1	8.5	

(أ) هل تدل هذه البيانات على أن هناك اختلاف في متوسط ذوبان الأوكسجين بين المواقع الأربعة (عند مستوى معنوية $\alpha=0.05$).

(ت) استعمل اختبار دنكن للمقارنة بين المتوسطات للمواقع المختلفة ؟

-٣٥- قام باحث في مجال الزراعة بدراسة لمقارنة معدل النمو لنبات مائي في أربعة مواقع. جزء من الدراسة تناول طول الورقة لهذا النبات اختبرت عينات عشوائية من كل موقع. تعطي البيانات التالية متوسط طول الورقة لكل نبات (مقاسه بالسنتيمتر) لعينة عشوائية من 10 ورقات لكل نبات .

	1	5.7	6.3	6.1	6.0	5.8	6.2
	2	6.2	5.3	5.7	6.0	5.2	5.5
الموقع	3	5.4	5.0	6.0	5.6	4.9	5.2
	4	3.7	3.2	3.9	4	3.5	3.6

هل توجد فروق معنوية بين المواقع عند مستوى معنوية $\alpha=0.05$ ؟

-٣٦- بفرض أن خمسة أنواع من الخلطات diets غُذيت بها خمسة مجموعات من الفئران متشابهين تماماً وسُجلت الزيادة في الوزن والنتائج في الجدول التالي

الخلطة	n_i	\bar{x}_i	s_i^2
Red , Maple '74	13	1.134	0.0252
Red , Oak / red maple	10	1.148	0.253
Red Maple '75	20	1.159	0.0179
Red Oak	16	1.191	0.0200
Red Oak/white pine	16	1.217	0.016

أ- أحسب $\bar{x}_{..}$.

ب- المطلوب تحليل هذه البيانات بأسلوب تحليل التباين؟

-٣٧- قام باحث بعمل مقارنة لستة أنواع من الزيت الصناعي تخص الأحماض الدهنية الغير مشبعة والبيانات في الجدول التالي:

Imperial	14.1	13.6	14.4	14.3	
Parkay	12.8	12.5	13.4	13	12.3
Blue Bounet	13.5	12.7	12.6	13.9	
Chiffon	16.8	17.2	16.4	17.3	
Mazola	16.8	17.2	16.4	17.3	18.0
Fleischmann's	18.1	17.2	18.7	18.4	

المطلوب تحليل هذه البيانات بأسلوب تحليل التباين ؟

٣٨- أجرى باحث في مجال العلوم الميكروبيولوجي اختبار الظاهرة الضاعف الضوئي photoreactive phenomenon في البكتريا عند تعريضها للأشعة : (1) العينة الأولى تم تعريضها للضوء المرئي لمدة 10 دقائق قبل وضعها في الحضانة (2) العينة الثانية تم تعريضها للضوء المرئي لمدة 5 دقائق قبل وضعها في الحضانة (3) العينة الثالثة وضعت في الحضانة بدون تعريضها للضوء البيانات التالية تمثل عدد مستعمرات البكتريا وذلك خمسة تكرارات لكل معالجة :

- (1) 90, 100, 75, 70, 64
(2) 72, 81, 53, 48, 55
(3) 45, 40, 10, 23, 32

المطلوب :

(أ) اكتب جدول تحليل التباين

(ب) اختبر معنوية الفروق بين المعالجات المختلفة

(ث) في حالة ما إذا كانت الفروق معنوية أى من هذه المعالجات يختلف عن الآخر؟

٣٩- بالرغم من أن الشاي يعتبر من أكثر المشروبات انتشارا بعد الماء إلا أن الكثير لا يعلم عن قيمته الغذائية nutritional value . فهناك folocin وهو واحد من مجموعة فيتامين B والذي يوجد بكميات معنوية في الشاي . استخدمت طرق تحليل دقيقة لحساب كمية folocin في الشاي في عينات عشوائية اختبرت من أربعة أنواع من الشاي الأخضر . البيانات في الجدول التالي :

نوع	1	7.9	6.2	6.6	8.6	8.9	10.1	9.6
	2	5.7	7.5	9.8	6.1	8.4		
الشاي	3	6.8	7.5	5	5	5.3	6.1	
	4	4.6	7.1	7.9	7.9	5.4		

هل البيانات السابقة تدل على أن متوسط folocin واحد في كل أنواع الشاي ؟ (وذلك عند مستوى معنوية $\alpha=0.01$).

٤٠- تغطي البيانات التالية محصول الطماطم (kg/plot) أثناء الزراعة بالمعروض لأربعة مستويات من التوصيل الكهربائي electrical conductivity .

المستويات	1.6	59.5	53.3	56.8	63.1
	3.8	55.2	59.1	52.8	54.9
	6.0	52.7	48.8	53.9	49
	10.2	44.6	48.5	41	47.3

والمطلوب اختبار معنوية الفروق بين متوسطات المستويات عند مستوى معنوية $\alpha=0.01$.

٤١- يعتقد باحث أن درجة الحرارة والملوحة عاملان مهمان في إنتاجية محصول الجميري . لذلك تم تصميم تجربة ذات عاملين ، العامل الأول (الملوحة) له ثلاثة مستويات والعامل الثاني درجة الحرارة وله ثلاثة مستويات . وقد تم تربية الجميري في المستويات السابقة من الملوحة والحرارة . وتسجيل محصول الجميري في الجدول التالي (البيانات مقاسه بعدد التانكات الكبيرة من سعة 80 جالون) .

الحرارة	الملوحة		
	A	B	C
60°F	3	5	4
70°F	11	10	12
80°F	16	2	17

أ- أوجد جدول تحليل التباين لهذه التجربة ؟

ب- هل تبين البيانات أن هناك فرقاً معنوياً بين متوسطات المعالجات عند مستوى معنوية $\alpha=0.01$.

٤٢- تعطي البيانات في الجدول التالي مبيعات منظم لعلالين وثلاثة تركيبات مختلفة والمطلوب الإجابة على السالزات الآتية :

(أ) هل يختلف متوسط المبيعات باختلاف تركيبة الصابون ؟

(ب) هل يختلف متوسط المبيعات باختلاف نوع العلاف ؟

العلاف	التركيبة		
	تركيبة (1)	تركيبة (2)	تركيبة (3)
غلاف 1	83	75	79
غلافه 2	74	75	78

٤٣- في دراسة على تأثير عادم السيارات على تلوث الهواء أخذت عينات من الهواء عند أزمة مختلفة وعند مواقع مختلفة وتم تحليلها لمعرفة كمية المادة المسببة للتلوث الموجود في الهواء مقاسه mg/m^3 . البيانات في الجدول التالي

الوقت	الأزمة				
	1	2	3	9	5
أكتوبر ١٩٧٢	76	67	81	56	51
يناير ١٩٧٦	82	69	96	59	70
مايو ١٩٧٦	68	59	67	54	42
سبتمبر ١٩٧٦	63	56	64	58	37

(أ) أختبر معنوية الفروق بين المواقع ؟

(ب) أختبر معنوية الفروق بين الأوراق ؟

-٤٤- في تجربة لمقارنة أربعة أنواع من الطلاء وثلاثة أنواع من البكرات (المستخدمة في الطلاء) استخدم جالون من كل نوع واستخدم بكرة للطلاء وتعطى البيانات التالية عدد الأقدام المربعة التي غطاها الطلاء .

نوع الدهان	البكرة			
	1	454	446	451
	2	446	444	442
	3	439	442	444
	4	443	437	443

(أ) أختبر معنوية الفروق بين البكرات ؟

(ب) أختبر معنوية الفروق بين الدهانات ؟

-٤٥- في دراسة لمقارنة ثلاثة مستويات من digitalis على مستوى الكالسيوم في عضلة قلب الكلب (الوصف الحقيقي للتجربة ثم حذفه) حيث أخذت أنسجة القلب لأربعة حيوانات كل نسيج وزع عشوائيا على المعالجات الثلاثة وقد تم تقدير مستوى الكالسيوم في الأنسجة للمعالجات الثلاثة والبيانات في الجدول التالي :

المعالجة	مستويات الأنسجة			
	1	2	3	4
A	1342	1387	1549	1150
B	1881	1140	1296	1579
C	1608	1698	1029	1319

(أ) أجري تحليل التباين . هل هناك تفاوت في متوسطات المعالجات المختلفة ؟ استعمل

مستوى معنوية $\alpha=0.05$.

(ب) هل هناك تفاوت في متوسطات العينات المختلفة (استعمل مستوى معنوية

$\alpha=0.05$.

-٤٦- فيما يلي درجات سمة الانبساطية لدى أربعة مجموعات وفي داخل كل مجموعة ذكور وإناث .

المجموعة	الجنس	
	ذكور	إناث
المجموعة الأولى	5, 4, 3, 2, 6	3, 2, 4, 5, 6

المجموعة الثانية	7, 5, 2, 4, 3	7, 4, 5, 3, 2
المجموعة الثالثة	6, 7, 8, 9, 10	8, 8, 9, 77
المجموعة الرابعة	9, 18, 8, 7, 16	15, 8, 9, 8, 8

(أ) المطلوب التحقق من صحة الفرض القائل .. لا توجد فروق معنوية في درجة السمة الانبساطية بين المجموعات المختلفة .

(ب) لا توجد فروق معنوية في سمة الانبساطية بين الذكور والإناث .

(ج) لا يوجد تفاعل بين الجنس و التجميع المختلفة .

٤٧- طبق اختبار للقلق على مجموعتين من الذكور والإناث في مجموعات عمرية مختلفة وجاءت درجاتهم كما يلي :

المراحل العمرية	الجنس	
	ذكور	إناث
الطفولة	2, 3, 2, 4, 5	3, 4, 4, 2, 7
المراهقة	13, 15, 12, 8, 11	12, 6, 17, 7, 12
الشباب	5, 6, 6, 7, 4	5, 4, 5, 6, 7, 7

المطلوب :

(أ) هل توجد فروق بين الجنسين في القلق ؟

(ب) هل توجد فروق بين فئات العمر في القلق ؟

(ج) هل يوجد تداخل بين عامل الجنس والعمر على القلق ؟

٤٨- أجريت دراسة طبية على تأثير ثلاثة أنواع من الأدوية على سلوك مجموعتين مسن المرضى النفسيين (إكتائيين - انفصامين) الدرجات التي حصلوا عليها في الجدول التالي :

نوع المرضى	نوع الدواء		
	1	2	3
إكتائيين	4, 8, 0	10, 8, 6	8, 6, 4
أنفصامين	4, 10, 6	4, 2, 0	15, 12, 9

(ب) أختبر معنوية الفروق بين الأدوية ؟

(ت) أختبر معنوية الفروق بين النوعين من المرضى ؟

(ث) هل هناك تفاعل بين نوع الدواء ونوع المرضى ؟

٤٩- تم إجراء تجربة زراعية للدراسة تأثير الأنواع المختلفة وكذلك طرق الزراعة المختلفة (من حيث كثافة النباتات في مساحة معينة وهي كالتالي 10,20, 30, 40 (الف نبات لكل هكتار) على إنتاجية محصول الطماطم. البيانات في الجدول التالي :

كثافة النباتات				
النوع	10,000	20,000	30,000	40,000
H	10.5, 9.2, 7.9	12.8, 11.2, 13.3	12.1, 12.6, 14	10.8, 9.1, 12.5
Ife	8.1, 8.6, 10.1	12.7, 13.7, 11.5	14.4, 15.4, 13.7	11.3, 12.5, 14.5
P	16.1, 15.3, 17.5	16.6, 19.9, 18.5	20.8, 18, 21	18.4, 18.9, 17

المطلوب تحليل هذه البيانات للفروق المعنوية في المتوسطات بين الأنواع المختلفة من الطماطم وبين مستويات الكثافات المختلفة والتفاعل بين النوع والكثافة للنباتات عند مستوى معنوية $\alpha=0.05$.

٥٠- في تجربة للدراسة تأثير الكميات المختلفة من Carbon fiber والكميات المختلفة من الرمل المضافة على عملية molding وذلك خلال صناعة الورقة ، تم الحصول على البيانات التالية :

الرمل المضاف	Carbon fiber		
	0.0	0.25	0.5
0.0	61.0 63.0	69 69	76 69
15	69 67	69 74	69 74
30	56 74	74 72	74 74

المطلوب تحليل هذه البيانات بأسلوب تحليل التباين عند مستوى معنوية $\alpha=0.01$.

٥١- في تجربة زراعة للدراسة تأثير الأنواع المختلفة من البطاطا A, B, C وكذلك المناطق الجغرافية المختلفة على إنتاجية محصول البطاطا ، تم اختبار 9 قطع متساوية في المساحة في كل منطقة جغرافية ثم زراعة كل صنف من الأصناف الثلاثة في 3 قطع اختبرت عشوائياً والبيانات في الجدول التالي :

المنطقة الجغرافية	أنواع البطاطا		
	A	B	C
1	14	19	21
	18	23	16
	11	17	13
2	16	23	18
	9	17	20
	12	21	21
3	8	11	9
	5	16	6
	12	9	7
4	4	20	18
	7	15	14
	10	13	11

أوجد جدول تحليل التباين ثم اختبر التأثيرات والتفاعلات عند مستوى $\alpha=0.05$.

٥٢- قام مهندس بقياس التيار (A) والضروري لإنتاج مستوى معين من الوضوح (brightness) بصمام التليفزيون وذلك لأنواع مختلفة من الزجاج وأنواع مختلفة من الفوسفور والبيانات في الجدول التالي :

		نوع الفسفور		
نوع الزجاج	1	1	2	3
	2	280, 290, 285	300, 310, 295	270, 285, 290
	3	230, 235, 240	260, 235, 240	220, 225, 230

والمطلوب تحليل هذه البيانات بأسلوب تحليل التباين .

٥٣- اختبرت خمس عينات عشوائية من الطحالب plankton من المناطق A, B, C على بحيرة خلال شهر مايو . وقد كررت التجربة مرة أخرى في شهر أغسطس . البيانات (مقاسه thousands of plankton) معطاة في الجدول التالي :

(أ) اختبر معنوية الفروق بين المواقع . ، (ب) اختبر معنوية الفروق بين أوقات الجمع . (ج) اختبر التفاعل بين الموقع ومعاد الجمع .

الموقع	معداد الجمع	
	أغسطس	مايو
A	97, 102, 109, 99, 101	107, 112, 118, 108, 111
B	105, 110, 115, 110	110, 115, 119, 110, 112

٥٤- تعطى البيانات التالية الحموضة الكلية لعينات من ثلاثة أنواع من القمح تم تحليلها باستخدام تركيزات مختلفة من مركب ethanoloic Na OH .

المركب	نوع القمح		
	Morwell	Yallourn	Maddingley
0.404 N	8.27, 8.17	8.66, 8.61	8.14, 7.96
0.626 N	8.03, 8.21	8.42, 8.28	8.02, 7.89
0.786 N	8.60, 8.20	8.61, 8.76	8.13, 8.07

أوجد جدول تحليل التباين واختبر التفاعل ثم اختبر التأثيرات الأساسية لكل عامل (عند مستوى معنوية 0.0).

٥٥- تعطى البيانات التالية نتائج أربعة اختبارات في أربعة مقررات في جامعة ما خمسة من الطلبة.

الطالب	اللغة الإنجليزية		اللغة الفرنسية		الإحياء		الرياضيات	
1	50	57	72	80	86	80	89	62
	71	64	76	77	90	77	78	79
2	89	94	81	35	80	92	78	95
	66	87	79	67	61	92	55	67
3	73	46	90	95	77	70	66	65
	58	81	59	91	83	72	50	88
4	75	48	42	51	55	48	85	59
	25	75	41	31	12	55	63	70
5	93	93	95	80	80	76	89	76
	62	75	97	95	86	80	86	94

أستخدم مستوى معنوية $\alpha=0.05$ لإختبار الفروض التالية :

(أ) هل هناك تفاوت في صعوبة المقررات الأربعة ؟

(ب) هل هناك تفاوت في مقدرات الطلبة . استعمل مستوى دلالة 0.05 .

(ت) هل يوجد تداخل بين الطلبة والمقررات ؟

٥٦- في تجربة لمقارنة أربعة أنواع من الإطارات وثلاثة أنواع من الطرق علسى عدد آلاف الكيلو مترات التي قطعت قبل تلف الإطار تم الحصول على البيانات التالية :

الطرق	نوع الإطار			
	A	B	C	D
1	9.9	8.9	7.5	9.4
	10.5	7.9	7.3	9.7
	11.3	7.6	7.2	10.6
	10.6	7.4	8.3	9.5

2	8.3	01.0	8.6	10.2
	7.2	9.9	8.4	10.3
	7.8	9.6	8.5	9.3
	7.1	10.5	9.2	9.2
3	9.6	8.6	6.8	9.2
	8.2	8.3	6.8	9.5
	8.1	8.5	7.3	8.6
	9.2	8.1	7.2	8.2

أستخدم مستوى معنوية $\alpha=0.01$ لاختبار الفروض التالية :

(أ) هل هناك فروق معنوية بين الأنواع ؟

(ب) هل هناك فروق معنوية بين الطرق ؟

(ث) هل يوجد تفاعل بين الأنواع والطرق ؟

٥٧- في دراسة لتقدير تأثير نوعين من الإعلانات على كمية المبيعات لثلاثة أنواع من الكيك تم تسجيل المبيعات لكل نوع بعد الإعلان أ ثم بعد الإعلانات ب . ثم كررت التجربة ثلاثة مرات لكل إعلان . نتائج التجربة في الجدول التالي

الإعلان			
		أ	ب
الكيك	A	571, 563, 559	1091, 1076, 1065
	B	553, 570, 550	1027, 1072, 999
	C	576, 453, 591	1065, 1077, 1051

(أ) أوجد جدول تحليل التباين ؟

(ب) أختبر معنوية الفروق بين الأنواع المختلفة عند مستوى معنوية

$$\alpha=0.05$$

(ج) أختبر معنوية الفروق بين الإعلانات عند مستوى معنوية $\alpha=0.05$.

(د) هل هناك تفاعل بين أنواع الكيك والإعلانات عند مستوى معنوية $\alpha=0.05$.

٥٨- أجرى اختبار على مجموعة من الأطفال وذلك لبيان تأثير الدافع الشخصي أو تشجيع الوالدين على ذكاء الأطفال وكانت درجات الذكاء كالآتي :

تشجيع الوالدين	دافع شخصي	
	عالي	منخفض
عالي	121, 116, 110, 103, 104, 115, 117, 116, 118, 119	112, 101, 111, 102, 79, 74, 91, 91, 80, 89

منخفض	116, 93, 94, 93, 98, 95, 98, 97, 95, 94	69, 65, 91, 91, 65, 70, 71, 72, 72, 73
-------	--	---

المطلوب تحليل هذه البيانات بأسلوب تحليل التباين عند مستوى معنوية $\alpha=0.05$.

الفصل الثاني عشر

الاختبارات الالعلميه

Nonparametric Tests

Introduction

(١٢-١) مقدمة

تعتمد الطرق المستخدمة في اختبارات الفروض وتحليل التباين وتحليل الانحدار وتحليل الارتباط (الطرق المعلمية parametric methods)، والتي سبق مناقشتها في الفصول السابقة، على عدد من الفروض . فعلي سبل المثال يشترط في الاختبار الذي يحس متوسط مجتمع (اختبار t) أن المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً وذلك عندما يكون تباين المجتمع غير معروف وحجم العينة صغير . أيضا في تحليل التباين يشترط أن المجتمعات التي اختيرت منها العينات تتبع توزيعات طبيعية وتبايناتها متساوية . الاختبارات السابقة سوف تكون غير مجدية إذا لم تتحقق الشروط الخاصة بها ، وفي هذه الحالة تكون الاختبارات اللامعلمية هي البديل . عموماً تستخدم الاختبارات اللامعلمية في الأحوال الآتية :

- (أ) عندما تكون الشروط اللازمة للاختبار المعلمي غير مستوفاة .
- (ب) عندما يدور فرض العلم والفرض البديل على أشياء وصفية وليست على معلمة مجتمع كما في اختبار جودة التوليد والذي سوف نتناوله في البند التالي . أيضا عند الرغبة في عمل مقارنة بين مجتمعين أو أكثر وذلك بالاعتماد على عينات عشوائية مختارة من هذه المجتمعات دون التعرف على التوزيعات الاحتمالية أو التعرض لها .
- (ج) عندما يكون مقياس البيانات وصفي أو ترتيبي .
- (د) عند الحاجة للوصول إلى قرار سريع بدون استخدام الآلات الحاسبة أو الحاسبات الإلكترونية .

بعض مميزات الطرق اللامعلمية

- (أ) فرض عدم الدقة قليلة عند تطبيقها وذلك لاعتمادها على أقل قدر من الفروض .
- (ب) لبعض الطرق اللامعلمية ، العمليات تتم بسرعة لذلك فإنها توفر الوقت وخصوصاً عند عدم توفر الآلات الحاسبة .
- (ج) تلائم كثير من الباحثين في مجال علم النفس والاجتماع لأن معظم البيانات الذين يعملون معها تكون من النوع الوصفي أو الترتيبي .
- (د) تلائم الباحثين الذين لديهم أدنى معلومات في مجال الرياضيات والإحصاء وذلك لسهولة المفاهيم والطرق اللامعلمية .

بعض عيوب الطرق اللامعلمية

- (أ) ولأن العمليات المستخدمة في معظم الاختبارات اللامعلمية بسيطة وسريعة فإنها تؤدي إلى فقد في المعلومات الموجودة في البيانات كما هو الحال عند تحويل البيانات إلى رتب ، وهذا يؤدي إلى فقد كبير في الدقة .
- (ب) بعض الطرق اللامعلمية تكون معقدة .
- (١٢-٢) اختبار مربع كاي لجودة التوفيق

Chi-square Goodness-of-fit Test

أن عملية التعرف على التوزيع الاحتمالي للمجتمع الذي اختيرت منه عينة عشوائية من الشروط اللازمة لتطبيق بعض الاختبارات كما أوضحنا في البند السابق . يستخدم اختبار مربع كاي لجودة التوفيق لاختبار ما إذا كانت مشاهدات عينة عشوائية تم اختيارها من مجتمع له توزيع احتمالي معين .

تكون البيانات اللازمة للاختبار من عينة عشوائية من الحجم n من المشاهدات المستقلة . تصنف المشاهدات إلى k من الفئات الشاملة الماتعة (الخلايا cell) كما هو موضح في جدول (١٢-١) . عدد المشاهدات التي تقع في فئة معطاة تسمى التكرار المشاهد observed frequency لهذه الفئة حيث O_i تمثل عدد المشاهدات في الفئة رقم i ، $i = 1, 2, \dots, k$.

جدول (١٢-١)

الفئة	1	2	...	i	...	k
التكرار	O_1	O_2	...	O_i	...	O_k
المشاهد						

الفئات قد تكون اسمية (وصفية) أو عددية . على سبيل المثال قد تنتمي مشاهدات العينة إلى واحدة من الفئتين الاسمين ذكر وأنثى . أيضا إذا كان المتغير موضع الدراسة هو العمر فإن مشاهدات العينة قد تتبع واحدة من الفئات العمرية التالية :

$$<15, 15-24, 25-34, 35-44, 45-54, \geq 55$$

لإجراء الاختبار نعرف الاحتمال ، سوف يرمز له بالرمز P_i ، بأن مشاهدة اختيرت عشوائياً من المجتمع النظري (المفترض) سوف تقع في الفئة رقم i وعلى ذلك نعين الاحتمالات P_1, P_2, \dots, P_k للفئات $i = 1, 2, \dots, k$ على التوالي . عندما يكون فرض العدم صحيح ، فإنه يمكن حساب التكرارات المتوقعة لكل فئة أي $n_1 P_1, n_2 P_2, \dots, n_k P_k$ للفئات $i = 1, 2, \dots, k$ على التوالي والتي سوف يرمز لها بالرمز E_1, E_2, \dots, E_k . فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : العينة اختيرت من مجتمع يتبع توزيع احتمالي معين .

H_1 : العينة اختيرت من مجتمع لا يتبع هذا التوزيع الاحتمالي المعين .

للعينات الكبيرة وبفرض أن H_0 صحيح فإن :

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

قيمة لمطير عشوائي X^2 تقريباً يتبع توزيع χ^2 بدرجات حرية $k-1$. مستوى معنوية α فإن منطقة الرفض $X^2 > \chi^2_{\alpha}$ حيث أن χ^2_{α} تستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $k-1$. إذا وقعت χ^2 في منطقة الرفض نرفض H_0 . يكون التقريب مقبولاً إذا كان عدد مشاهدات العينة أكبر من 50 والتكرار المتوقع المناظر لكل فئة لا يقل عن 5 . في بعض الأحيان تستخدم مشاهدات العينة في تقدير معلمة أو أكثر من معالم المجتمع ثم يستخدم هذا التقدير في حساب التكرارات المتوقعة . فإذا كان عدد المعالم المقدرة هو m فإن درجات الحرية في هذه الحالة تصبح $(k-m-1)$.

مثال (١٢-١) يقوم المستول في مصنع لإنتاج معاطف السيدات العالية الجودة بإرسال كل معطف منتج إلى واحد من مراكز مراقبة الجودة . يعطى جدول (١٢-٢) التوزيع التكراري لعدد المعاطف المرفوضة لعينة عشوائية من الحجم $n=100$. هل يمكن القول أن عدد المعاطف المرفوضة واحدة لكل مركز ؟

جدول (١٢-٢)

مركز الفحص	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2 / E_i$
1	15	10	5	25	2.5
2	3	10	-7	49	4.9
3	7	10	-3	9	0.9
4	8	10	-2	4	0.4
5	5	10	-5	25	2.5
6	12	10	+2	4	0.4
7	11	10	+1	1	0.1
8	13	10	+3	9	0.9
9	9	10	-1	1	0.1
10	17	10	+7	49	4.9
المجموع	100	100			17.6

الحل.

H_0 : المراقبون يرفضون نفس العدد من المعاطف (العينة تم اختيارها من مجتمع يتبع التوزيع المنتظم).

H_1 : المراقبون لا يرفضون نفس العدد من المعاطف (العينة تم اختيارها من مجتمع لا يتبع التوزيع المنتظم).

تحت فرض العدم فإن $P_i = \frac{1}{10}, i = 1, 2, \dots, 10$. التكرارات المتوقعة في جدول (١٢-٢) تم حسابها من الصيغة $E_i = nP_i = (100)(\frac{1}{10}) = 10$ حيث $i = 1, 2, \dots, 10$. مستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{0.05} = 16.919$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) عند درجات حرية $9 = 10 - 1 = k - 1$. منطقة الرفض $\chi^2 > 16.919$. وبما أن χ^2 تقع في منطقة الرفض نرفض H_0 .

مثال (١٢-٢) يعطى الجدول (١٢-٣) عدد الوفيات في الأسبوع x والتي وقعت في 10 مدن خلال 200 أسبوع والنتيجة من حوادث السيارات .

جدول (١٢-٣)

x	0	1	2	3	4
التكرارات	109	65	22	3	1
المشاهدة					

هل النتائج المعطاة في جدول (١٢-٣) تتفق مع الفرض القائل أن عدد الوفيات في الأسبوع والنتيجة من حوادث السيارات تتبع توزيع بواسون ؟ استخدم مستوى معنوية $\alpha = 0.05$.

الحل .

H_0 : عدد الوفيات في الأسبوع والنتيجة من حوادث السيارات تتبع توزيع بواسون .

H_1 : عدد الوفيات في الأسبوع والنتيجة من حوادث السيارات لا تتبع توزيع بواسون .

بفرض أن H_0 صحيح فإن الدالة الاحتمالية لتوزيع بواسون هي :

$$P(X = i) = \frac{\bar{e}^{\mu} \mu^i}{i!}, i = 1, 2, \dots$$

وحيث أن μ غير محددة من فرض العدم فإنه يمكن تقديرها من مشاهدات العينة وأفضل تقدير للمعلمة μ هو الوسيط الحسابي للملاحظات والذي يتم حسابه من جدول (٣-١٢) حيث أن :

$$\hat{\mu} = \bar{x} = \frac{\sum x_i O_i}{\sum O_i} = \frac{122}{200} = 0.61.$$

باستبدال μ في توزيع بواسون بالقيمة $\hat{\mu}$ نحصل على :

$$P_i = P(X=i) = \frac{(0.61)^i \bar{e}^{0.61}}{i!}, i=0,1,2,3,4.$$

التكرارات المتوقعة لعدد الوفيات في الأسبوع والناجمة من حوادث السيارات ،
 $E_i = nP_i, i=0,1,2,3,4$ ، معطاة في جدول (٤-١٢) .

جدول (٤-١٢)

x	التكرارات المتوقعة
0	$200.P(X=0) = 200 \frac{(0.61)^0 \bar{e}^{0.61}}{0!} = 108.7$
1	$200.P(X=1) = 200 \frac{(0.61)^1 \bar{e}^{0.61}}{1!} = 66.3$
2	$200.P(X=2) = 200 \frac{(0.61)^2 \bar{e}^{0.61}}{2!} = 20.2$
3	$200.P(X=3) = 200 \frac{(0.61)^3 \bar{e}^{0.61}}{3!} = 4.1$
4	$200.P(X=4) = 200 \frac{(0.61)^4 \bar{e}^{0.61}}{4!} = 0.7$

يعطي الجدول (٥-١٢) التكرارات المشاهدة والتكرارات المتوقعة بعد تقريبها .

جدول (٥-١٢)

x	0	1	2	3	4	5 أو أكثر	المجموع
التكرارات المتوقعة	109	66	20	4	1	0	200
التكرارات المشاهدة	109	65	22	3	1	0	200

بالنظر إلى التكرارات المتوقعة في جدول (١٢-٥) نجد أن هناك تكرارين كل منهما أقل من 5 وهما التكرار الرابع والتكرار الخامس وبذلك لا يصلح تطبيق اختبار مربع كاي إلا بعد التغلب على هذه المشكلة وذلك بدمج التكرارين كما هو معطى في جدول (١٢-٦) :

جدول (١٢-٦)

x	0	1	2	أكبر أو يساوى 3
التكرارات المتوقعة	109	66	20	5
التكرارات المشاهدة	109	65	22	4

من جدول (١٢-٦) فإن :

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} = 0.41515.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi_{0.05}^2 = 5.992$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $k-m-1=4-1-1=2$. منطقة الرفض $\chi^2 > 5.992$. وبما أن χ^2 تقع في منطقة القبول نقبل H_0 .

مثال (١٢-٣) يعطى جدول (١٢-٧) التوزيع التكراري لأطوال 40 بطارية. هل البيانات في جدول (١٢-٧) تتفق مع القول أن التوزيع الطبيعي يعطى توفيق جيد لتوزيع أعمار البطاريات وذلك عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢-٧)

الحدود الفعلية للفئة	التكرارات المشاهدة	التكرارات المتوقعة
1.45 - 1.95	2	0.6
1.95 - 2.45	1	2.6
2.45 - 2.95	4	6.8
2.95 - 3.45	15	10.7
3.45 - 3.95	10	10.3
3.95 - 4.45	5	6.1
4.45 - 4.95	3	2.2
المجموع	40	

الحل .

H_0 : العينة اختيرت من مجتمع يتبع التوزيع الطبيعي .

H_1 : العينة اختيرت من مجتمع لا يتبع التوزيع الطبيعي .

التكرارات المتوقعة في جدول (٧-١٢) تم حسابها من منحني طبيعي له نفس المتوسط والانحراف المعياري لملاحظات العينة المطبوعة في جدول (٧-١٢) حيث $\bar{x} = 3.4125$, $s = 0.697$ أي أننا سوف نستخدم \bar{x} , s (بدلاً من μ , σ الخاصة بالمتجمع الذي اختيرت منه العينة) في حساب قيم z على سبيل المثال للفئة الرابعة فإن :

$$z_1 = \frac{2.95 - 3.4125}{0.697} = -0.67,$$

$$z_2 = \frac{3.45 - 3.4125}{0.697} = 0.054.$$

ومن جدول التوزيع الطبيعي القياسي في ملحق (٣) فإن المساحة بين $z_1 = -0.67$, $z_2 = 0.054$ هي :

$$\begin{aligned} P(-0.67 < Z < 0.054) \\ &= P(0 < Z < 0.67) + P(0 < Z < 0.054) \\ &= 0.2486 + 0.0199 = 0.2685. \end{aligned}$$

وعلى ذلك فإن التكرارات المتوقعة للفئة الرابعة هي :

$$E_4 = (0.2685)(40) = 10.7$$

حيث أن $\sum_{i=1}^k O_i = 40$ التكرار المتوقع للفئة الأولى تم الحصول عليه باستخدام المساحة الكلية

تحت المنحني الطبيعي على يسار القيمة 1.95 (الحُد الأعلى الفعلي للفئة الأولى). للفئة الأخيرة استخدمت المساحة الكلية تحت المنحني الطبيعي على يمين القيمة 4.45 (الحُد الأدنى الفعلي للفئة الأخيرة). التكرارات المتبقية تم حسابها بنفس الطريقة التي شرحناها للفئة الرابعة. بدمج التكرارات المتوقعة للفئات التي تكراراتها المتوقعة أقل من 5 نحصل على جدول (٨-١٢) .

جدول (٨-١٢)

الحدود الفعلية	التكرارات المشاهدة	التكرارات المتوقعة
1.45 - 2.95	7	10
2.95 - 3.45	15	10.7
3.45 - 3.95	10	10.3
3.95 - 4.95	8	8.3

من جدول (٨-١٢) فإن :

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} = 2.648.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{0.05} = 3.843$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $k-m-1=4-2-1=1$. منطقة الرفض $\chi^2 > 3.843$. وبما أن χ^2 تقع في منطقة القبول نقبل H_0 .

(١٢-٣) اختبار مربع كاي للاستقلال The Chi-square Test of Independent

في كثير من الأحيان يرغب الباحث في التعرف عما إذا كانت هناك علاقة بين صفتين من صفات مجتمع ما . فعل سبيل المثال قد يرغب مسئول التغذية في مدرسة ما في التعرف عما إذا كانت الحالة الغذائية للطلاب لها علاقة بكفاءته التعليمية . أيضا قد يرغب باحث في مجال الوراثة في التعرف عما إذا كانت هناك علاقة بين لون الشعر ولون العينين ... الخ .

لإجراء الاختبار نختار عينة عشوائية من الحجم n من المجتمع موضوع الدراسة . تصنف مشاهدات هذه العينة حسب مستويات كل من الصفتين موضوع الدراسة في جدول مزدوج يسمى جدول التوافق Contingency table . بفرض أن A_1, A_2, \dots, A_k ترمز لمستويات الصفة A و B_1, B_2, \dots, B_k ترمز لمستويات الصفة B فإن جدول التوافق يكون على الشكل الموضح في جدول (١٢-٩) ، حيث أن O_{ij} ترمز لعدد المشاهدات التي يتوفر فيها المستوى A_i من الصفة A و المستوى B_j من الصفة B حيث $i = 1, 2, \dots, r$ و $j = 1, 2, \dots, c$. أيضا $n_{i.}$ ترمز لعدد المشاهدات التي يتوفر فيها المستوى A_i من الصفة A أي أن

$$n_{i.} = \sum_{j=1}^c O_{ij} \quad \text{أي أن } n_{.j} = \sum_{i=1}^r O_{ij} \quad \text{وعلى ذلك :}$$

$$\sum_{j=1}^c n_{.j} = \sum_{i=1}^r n_{i.} = \sum_{j=1}^c \sum_{i=1}^r O_{ij}.$$

يحتوي جدول التوافق على خانات (خلايا) عددها $(r \times c)$ خلية.

جدول (١٢-٩)

	B_1	B_2	...	B_c	المجموع
A_1	O_{11}	O_{12}	...	O_{1c}	$n_{1.}$
A_2	O_{21}	O_{22}	...	O_{2c}	$n_{2.}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
A_r	O_{r1}	O_{r2}	...	O_{rc}	$n_{r.}$
	$n_{.1}$	$n_{.2}$...	$n_{.c}$	n

فرض العدم والفرض البديل سوف يكونان علي الشكل :

H_0 : المتغيرين مستقلين .

H_1 : المتغيرين غير مستقلين .

يعتمد اختبار مربع كاي للاستقلال علي مقارنة التكرارات المشاهدة بالتكرارات المتوقعة في كل خلية عندما H_0 صحيح . إذا كان $P(A_i)$ يرمز لاحتمال أن يتوفر لمشاهدة ما المستوى A_i من الصفة A وإذا كان $P(B_j)$ يرمز لاحتمال أن يتوفر لمشاهدة ما المستوى B_j من الصفة B وإذا كان P_{ij} يرمز لاحتمال أن يتوفر لمشاهدة ما المستوى A_i من الصفة A و المستوى B_j من الصفة B فإن :

$$P_{ij} = P(A_i \cap B_j)$$

ولي حالة الاستقلال بين الصفتين A و B (تحت فرض العدم) فإن :

$$P_{ij} = P(A_i) \cdot P(B_j).$$

يمكن الحصول علي تقدير للاحتمال P_{ij} كالتالي :

$$P'_{ij} = \left(\frac{n_{i.}}{n} \right) \left(\frac{n_{.j}}{n} \right)$$

وعلي ذلك يمكن حساب التكرارات المتوقعة كالتالي :

$$E_{ij} = n \left(\frac{n_{i.}}{n} \right) \left(\frac{n_{.j}}{n} \right)$$

$$= \frac{n_{i.} \cdot n_{.j}}{n}, \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, c.$$

بافتراض أن H_0 صحيح فإن :

$$\chi^2 = \sum_{j=1}^c \sum_{i=1}^r \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

قيمة لمغير عشوائي X^2 تقريبا يتبع توزيع χ^2 بدرجات حرية $(r-1)(c-1)$ حيث r عدد الصفوف و c عدد الأعمدة في جدول التوافيق . لمستوى معنوية α فإن منطقة الرفض $X^2 > \chi^2_{\alpha}$ حيث χ^2_{α} تستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $(r-1)(c-1)$. إذا وقعت χ^2 في منطقة الرفض نرفض H_0 .

مثال (١٢ - ٤) يعتقد الأطباء أن عدد ساعات النوم لسيدة لديها أطفال يختلف عن عدد ساعات النوم قبل إنجابها . بفرض إنه تم سؤال 60 سيدة لديها أطفال وسجلت البيانات في جدول

(١٠-١٢) . ما هو الاستدلال الذي يمكن الحصول عليه من هذه البيانات ؟ عند مستوى معنوية $\alpha = 0.05$.

جدول (١٠-١٢)

عدد الأطفال	النوم الحالي بالمقارنة قبل الإنجاب			المجموع
	أقل	نفسه	أحسن	
1	25	5	0	30
2	10	4	1	15
3 أو أكثر	5	7	3	15
المجموع	40	16	4	60

الحل .

H_0 : عدد ساعات النوم وعدد الأطفال مستقلين .

H_1 : عدد ساعات النوم وعدد الأطفال غير مستقلين .

التكرارات المتوقعة معطاة في جدول (١١-١٢) .

جدول (١١-١٢)

عدد الأطفال	النوم الحالي بالمقارنة قبل الإنجاب		
	أقل	نفسه	أحسن
1	20	8	2
2	10	4	1
3 أو أكثر	10	4	1

التكرارات المشاهدة والتكرارات المتوقعة بعد دمج بعض التكرارات معطاة في جدول (١٢-١٢)

(١٢) و جدول (١٣-١٢) على التوالي وذلك حتى يتحقق الشرط أن عدد التكرارات المتوقعة

في كل خلية لا يقل عن 5 .

جدول (١٢-١٢)

عدد الأطفال	أقل	نفسه أو أحسن
1	25	$5 + 0 = 5$
2	10	$4 + 1 = 5$
3 أو أكثر	5	$7 + 3 = 10$

من جدول (١٢-١٢) و جدول (١٣-١٢) يمكن حساب :

$$\chi^2 = \sum_{j=1}^c \sum_{i=1}^r \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 11.25.$$

لمستوى معنوية $\alpha = 0.05$ فإن $\chi^2_{0.05} = 5.992$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $2 \times 1 = 2$. منطقة الرفض $X^2 > 5.992$. وبما أن χ^2 تقع في منطقة الرفض فإننا نرفض H_0 .

جدول (١٢-١٣)

عدد الأطفال	أقل	نفسه أو أحسن
1	20	10
2	10	5
3 أو أكثر	10	5

مثال (١٢-٥) : لدراسة العلاقة بين لون شعر الزوج والزوجة قام باحث باختيار عينة عشوائية من الحجم 500 (زوج وزوجة) وتم سؤالهم والبيانات في جدول (١٢-١٤).

جدول (١٢-١٤)

الزوجة	الزوج				المجموع
	أحمر	أصفر	أسود	بني	
أحمر	10	10	10	20	50
أصفر	10	40	50	50	150
أسود	13	25	60	52	150
بني	17	25	30	78	150
المجموع	50	100	150	200	500

المطلوب اختبار ما إذا كانت هناك علاقة بين لون شعر الزوج والزوجة أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل .

H_0 : لون شعر الزوج ولون شعر الزوجة مستقلين.

H_1 : لون شعر الزوج ولون شعر الزوجة غير مستقلين.

التكرارات المتوقعة معطاة في جدول (١٢-١٥) . من جدول (١٢-١٤) و جدول (١٢-١١) :
 ١٥) فإن :

$$\chi^2 = \sum_{j=1}^c \sum_{i=1}^r \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= 32.56.$$

جدول (١٢-١٥)

الزوجة	الزوج				المجموع
	أحمر	أصفر	أسود	بني	
أحمر	5	10	15	20	50
أصفر	15	30	45	60	150
أسود	15	30	45	60	150
بني	15	30	45	60	150
المجموع	50	100	150	200	500

لمستوى معنوية $\alpha = 0.05$ فإن $\chi_{0.05}^2 = 16.919$ والمستخرجة من جداول توزيع χ^2 في ملحق (٥) بدرجات حرية $3 \times 3 = 9$. منطقة الرفض $X^2 > 16.919$ وبما أن χ^2 المحسوبة تقع في منطقة الرفض فإننا نرفض H_0 .

إذا كان لكل من الصفين A, B مستويان فقط فإن الجدول الناتج يتكون من صفين وعمودين (أى أربع خلايا) . يسمى الجدول الناتج جدول الاقتران (2×2) . جدول (١٢-١٦) يمثل جدول القران . عدد درجات الحرية التي ترتبط بجدول الاقتران سوف تساوى الواحد الصحيح .

جدول (١٢-١٦)

الصفة الأولى	الصفة الثانية		
	B_1	B_2	
A_1	a	b	a + b
A_2	c	d	c + d
	a+c	b+d	n

يمكن استخدام صيغة بسيطة لحساب قيمة χ^2 كالتالي :

$$\chi^2 = \frac{n(ad - bc)^2}{(a+c)(b+d)(c+d)(a+b)}.$$

مثال (١٢-٦) لدراسة العلاقة بين النوم ليلا والتدخين اختبرت عينة عشوائية من 56 شخصا والبيانات معطاة في جدول (١٢-١٧) .

جدول (١٢-١٧)

التدخين	النوم		المجموع
	نعم	لا	
نعم	20	16	36
لا	6	14	20
المجموع	26	30	56

المطلوب اختبار ما إذا كانت هناك علاقة بين النوم ليلا والتدخين وذلك عند مستوى معنوية

$$\alpha = 0.05 .$$

الحل .

H_0 : المتغيرين مستقلين .

H_1 : المتغيرين غير مستقلين .

من جدول (١٢-١٧) فإن :

$$\chi^2 = \frac{n(ad - bc)^2}{(a+c)(b+d)(c+d)(a+b)}$$

$$= \frac{56[(20)(14) - (16)(6)]^2}{(26)(30)(20)(36)} = 3.376 .$$

لمسوى معنوية $\alpha = 0.05$ فإن $\chi^2_{0.05} = 3.843$ والمستخرجة من جدول توزيع χ^2 في ملحق

(٥) بدرجات حرية واحدة . منطقة الرفض $\chi^2 > 3.843$. وبما أن χ^2 تقع في منطقة

القبول نقبل H_0 .

سبق أن ذكرنا أن التكرارات المتوقعة في كل خلية يجب أن لا يقل عن 5 وإذا حدث وكان

أحد التكرارات المتوقعة أقل من 5 فإننا نقوم بدمج التكرارات . وعلي أي حال فإن هذه

الطريقة لا تستخدم في حالة جدول الاقتران . وقد اقترح Yates (1934) تصحيحا يستخدم في

حالة ما إذا كان أحد التكرارات المتوقعة أقل من 5 . وباستخدام التصحيح يصبح قيمة الإحصاء الذي يعتمد عليه قرارنا هو :

$$\chi^2 = \frac{n(|ad - bc| - \frac{n}{2})^2}{(a+c)(b+d)(c+d)(a+b)}$$

بتطبيق تصحيح Yates على البيانات في جدول (١٢-١٧) فإن قيمة الإحصاء تصبح :

$$\chi^2 = \frac{56[(20)(14) - (16)(6)] - \frac{56}{2}]^2}{(26)(30)(20)(36)} = 2.427.$$

لمستوى معنوية $\alpha = 0.05$ فإننا نحصل إلى نفس الاستنتاج الذي حصلنا عليه بدون تصحيح ، أى إننا نقبل H_0 .

(١٢-٤) اختبار مربع كاي للتجانس

The Chi-square Test of Homogeneity

نفرض أن لدينا مجتمعات عددها r وجميعها متماثلة من حيث التصنيف ونفرض أن c هي عدد فئات التصنيف في كل مجتمع . نفرض أن $P_{j|i}$ يرمز لنسبة مشاهدات المجتمع رقم i السقي تقع في الفئة رقم j . يمكن تمثيل هذه المجتمعات بالجدول (١٢-١٨) .

جدول (١٢-١٨)

المجتمع	فئات التصنيف						
	1	2	...	j	...	c	
1	$P_{1 1}$	$P_{2 1}$...	$P_{j 1}$...	$P_{c 1}$	1
2	$P_{1 2}$	$P_{2 2}$...	$P_{j 2}$...	$P_{c 2}$	1
⋮							⋮
i	$P_{1 i}$	$P_{2 i}$...	$P_{j i}$...	$P_{c i}$	1
⋮							⋮
r	$P_{1 r}$	$P_{2 r}$...	$P_{j r}$...	$P_{c r}$	1
	P_1	P_2	P_j	...	P_c	1

عندما تكون النسب $P_{j|i}$ مجهولة فإننا نرغب في معرفة ما إذا كانت المجتمعات التي عددها r متجانسة أي إننا نرغب في اختبار فرض العدم :

$$H_0 : P_{j|1} = P_{j|2} = \dots = P_{j|r} = P_j$$

$$; j = 1, 2, \dots, c.$$

لإجراء الاختبار فإننا نختار عينات عشوائية عددها r واحدة من كل مجتمع وأحجامها هي n_1, n_2, \dots, n_r على أن تكون العينات العشوائية مستقلة عن بعضها البعض. بفحص مشاهدات هذه العينات ووضع كل مشاهدة حسب تصنيفها نحصل على الجدول (١٢-١٩).

جدول (١٢ - ١٩)

المجتمع	1	2	...	j	...	c	
1	O_{11}	O_{12}	...	O_{1j}	...	O_{1c}	n_1
2	O_{21}	O_{22}	...	O_{2j}	...	O_{2c}	n_2
⋮							⋮
i	O_{i1}	O_{i2}	...	O_{ij}		O_{ic}	n_i
⋮							⋮
r	O_{r1}	O_{r2}	...	O_{rj}	...	O_{rc}	n_r
المجموع	$n_{.1}$	$n_{.2}$...	$n_{.j}$...	$n_{.c}$	n

$$\text{حيث } n = \sum_{i=1}^r n_i = \sum_{j=1}^c n_{.j} \text{ و } n_{.j} = \sum_{i=1}^r O_{ij} \text{ و } n_i = \sum_{j=1}^c O_{ij}$$

إذا كان فرض العدم صحيح وحيث أن النسب مجهولة فإننا نقوم بتقدير P_j حيث أن :

$$P'_j = \frac{n_{.j}}{n},$$

$$E_{ij} = n_i \cdot P'_j = n_i \left(\frac{n_{.j}}{n} \right),$$

وعلى ذلك فإن :

$$\chi^2 = \sum_{j=1}^c \sum_{i=1}^r \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

هي قيمة للمتغير العشوائي X^2 الذي تقريبا يتبع توزيع χ^2 بدرجات حرية $(r-1)(c-1)$. لاختبار فرض العدم عند مستوى معنوية α تتبع الخطوات التي استخدمناها في اختبار مربع كاي للاستقلال .

مثال (١٢-٧) قامت شركة للمياه الغازية بدراسة لمعرفة ما إذا كان هناك اختلاف بين شرائح مختلفة من المجتمع من ناحية التفضيل لثلاثة أنواع من المشروبات . استخدمت لهذه الدراسة أربع عينات مستقلة والنتائج معطاة في جدول (١٢-٢٠) . استخدم اختبار مربع كاي للتجانس لاختبار فرض العدم :

H_0 : التجمعات الأربعة متساوين في تفضيل المشروب.

ضد الفرض البديل :

H_1 : التجمعات الأربعة غير متساوين في تفضيل المشروب.

وذلك عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢-٢٠)

العينات الأربعة	نوع المشروب			المجموع
	A	B	C	
ربات البيوت	75	20	5	100
رجال الأعمال	50	130	20	200
عمال	5	25	17	47
طلبة	100	100	100	300
المجموع	230	275	142	647

الحل . التكرارات المتوقعة تم حسابها في جدول (١٢-٢١) .

جدول (١٢-٢١)

العينات الأربعة	نوع المشروب			المجموع
	A	B	C	
ربات البيوت	35.55	42.50	21.95	100
رجال الأعمال	71.10	85.01	43.89	200
عمال	16.71	19.98	10.32	47.01
طلبة	106.65	127.51	65.84	300
المجموع	230.01	275	142	647.01

نحسب قيمة الاحصاء من الصيغة التالية :

$$\chi^2 = \sum_{j=1}^c \sum_{i=1}^r \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$= 149.72.$$

لستوى معنوية $\alpha = 0.05$ فإن $\chi_{0.05}^2 = 12.592$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) عند درجات حرية $6 = 3 \times 2 = r \times c$. منطقة الرفض $X^2 > 12.592$. وبما أن χ^2 تقع في منطقة الرفض نرفض H_0 .

(١٢-٥) اختبار خاص بالاعتدال A Special Test for Normality

كثير من اختبارات الفروض التي تناولناها في الفصول السابقة تشترط أن مشاهدات العينة يجب أن تكون مختارة من مجتمع يتبع توزيعاً طبيعياً . فرض العدم والفرض البديل لا اختبار الاعتدال يكونان على الشكل :

H_0 : توزيع المجتمع الذي اختيرت منه العينة يتبع توزيعاً طبيعياً .

H_1 : توزيع المجتمع الذي اختيرت من العينة لا يتبع توزيعاً طبيعياً .

لإجراء الاختبار نختار عينة عشوائية من الحجم n من المشاهدات x_1, x_2, \dots, x_n وذلك من المجتمع موضع الدراسة . بفرض أن مشاهدات العينة تم ترتيبها من الأصغر إلى الأكبر (ترتيب تصاعدياً) حيث $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ ترمز إلى المشاهدات بعد ترتيبها . يتم حساب القيم $z_{(1)}, z_{(2)}, \dots, z_{(n)}$ من جدول التوزيع الطبيعي القياسي في ملحق (٣) والتي المساحة قبلها تساوي $[(i - 0.375)/(n + 0.25)]$. على سبيل المثال إذا كان حجم العينة $n = 20$ فإن $z_{(2)}$ تستخرج من جدول التوزيع الطبيعي القياسي في ملحق (٣) وهي القيمة التي المساحة قبلها $[(2 - 0.375)/(20 + 0.25)] = 0.0802$. أي أن $z_{(2)} \approx -1.4$. لأزواج القيم $(x_{(1)}, z_{(1)}), (x_{(2)}, z_{(2)}), \dots, (x_{(n)}, z_{(n)})$ يتم حساب معامل الارتباط البسيط r . بفرض أن H_0 صحيح فإن r قيمة لمطير عشوائي R_r له توزيع احتمالي . القيم الحرجة c_α من R_r معطاة في الجدول في ملحق (١٠) عند مستويات معنوية مختلفة . منطقة الرفض $R_r \leq c_\alpha$. إذا وقعت r في منطقة الرفض نرفض H_0 .

مثال (١٢-٨) يعطى الجدول (١٢-٢٢) مشاهدات لعينة عشوائية من الحجم $n=20$. كل قيمة تمثل نسبة العرض إلى الطول لقطعة مربعة من جلد الثعالب المستخدمة في صناعة حقائب للسيدات .

جدول (١٢-٢٢)

i	1	2	3	4	5	6	7	8	9	10
$x_{(i)}$	0.553	0.570	0.576	0.601	0.606	0.606	0.609	0.611	0.615	0.628
$z_{(i)}$	-1.87	-1.40	-1.13	-0.92	-0.66	-0.66	-0.45	-0.31	-0.19	-0.06
i	11	12	13	14	15	16	17	18	19	20
$x_{(i)}$	0.654	0.662	0.668	0.670	0.672	0.690	0.693	0.749	0.844	0.933
$z_{(i)}$	0.06	0.19	0.31	0.45	0.59	0.74	0.92	1.13	1.40	1.87

المطلوب اختبار فرض العدم :

 H_0 : توزيع المجتمع الذي اختبرت منه العينة طبيعي .

ضد الفرض البديل :

 H_1 : توزيع المجتمع الذي اختبرت منه العينة غير طبيعي .وذلك عند مستوى معنوية $\alpha = 0.01$.

الحل . من البيانات في جدول (١٢-٢٢) فإن :

$$n = 20 , \quad \Sigma x_{(i)} = 13.21 , \quad \Sigma x_{(i)}^2 = 8.887812 ,$$

$$\Sigma z_{(i)} = 0.01 , \quad \Sigma z_{(i)}^2 = 17.6039 , \quad \Sigma x_{(i)}z_{(i)} = 1.5371$$

وعلى ذلك :

$$r = \frac{\Sigma x_{(i)}z_{(i)} - \frac{\Sigma x_{(i)}\Sigma z_{(i)}}{n}}{\sqrt{\left[\Sigma x_{(i)}^2 - \frac{(\Sigma x_{(i)})^2}{n} \right] \left[\Sigma z_{(i)}^2 - \frac{(\Sigma z_{(i)})^2}{n} \right]}}$$

$$= \frac{1.5371 - \frac{(13.21)(0.01)}{20}}{\sqrt{\left[8.887812 - \frac{(13.21)^2}{20} \right] \left[17.6039 - \frac{(0.01)^2}{20} \right]}}$$

$$= 0.905.$$

لمستوى معنوية $\alpha = 0.01$ فإن $c_{res} = 0.929$ والمستخرجة من الجدول في ملحق (١٠) عند n . منطقة الرفض $R_r < 0.929$. بما أن r تقع في منطقة الرفض نرفض H_0 .

(١٢-٦) اختبار الإشارة لعينة واحدة

The One – sample Sign Testيستخدم اختبار الإشارة كدليل لاختبار t الخاص بمتوسط مجتمع وذلك عند عدم التحقق

من أن المجتمع الذي اختبرت منه العينة يتبع توزيعاً طبيعياً وحجم العينة أقل من 30 .

تكون البيانات اللازمة للتحليل من عينة عشوائية من الحجم n من المشاهدات x_1, x_2, \dots, x_n والمختارة من مجتمع متصل وسيطه مجهول . في الحقيقة إذا كان التوزيع متماثل فإن الوسيط يساوي الوسط الحسابي للمجتمع μ ويمكن استخدام اختبار الإشارة كاختبار للمتوسط . فرض العدم والفرض البديل سوف يكونان على الشكل :

$$H_0 : M = M_0,$$

$$H_0 : M > M_0.$$

لإجراء الاختبار نحسب القيمة k والتي تمثل عدد الإشارات السالبة للفروق $(x_i - M_0), i = 1, 2, \dots, n$ وإذا وجدت مشاهدة تساوي الوسيط نعملها ولا نأخذها في الاعتبار . بفرض أن H_0 صحيح فإن k تمثل قيمة لمتغير عشوائي (إحصاء) K له توزيع ذي الحدين بمعالم $p=0.5$, n . مستوى معنوية α نرفض H_0 إذا كان :

$$P(K \leq k | n, 0.5) \leq \alpha.$$

للفرض البديل $H_0 : M < M_0$ نرفض H_0 إذا كان :

$$P(K \leq k | n, 0.5) \leq \alpha.$$

حيث k تمثل عدد الإشارات الموجبة.

للفرض البديل $H_0 : M \neq M_0$ نرفض H_0 إذا كان :

$$P(K \leq k | n, 0.5) \leq \frac{\alpha}{2}$$

حيث k تمثل عدد الإشارات الموجبة أو السالبة أيهما اقل للفروق $(x_i - M_0), i = 1, 2, \dots, n$.

مثال (١٢ - ٩) يعطى الجدول (١٢ - ٢٣) الدخل السنوية (بالآلاف الدولارات) لعينة عشوائية من 21 عضو هيئة تدريس بإحدى الجامعات والمطلوب اختبار فرض العدم $H_0 : M = 25.1$: ضد الفرض البديل $H_1 : M > 25.1$ باستخدام مستوى معنوية $\alpha = 0.05$.

جدول (١٢ - ٢٣)

11.1	12.9	14.0	19.4	19.9	22.2	23.6	24.1	25.2	26.7	25.1
28.1	29.6	30.7	32.2	33.9	34.8	38.9	40.1	50.5	50.6	

لحساب عدد الإشارات السالبة k نحدد إشارات الفروق بين مشاهدات العينة والوسط حيث وجدت كالآتي :

$$-----++0++++++$$

وحيث أن لدينا مشاهدة تساوي صفر وبعد استبعادها يصبح حجم العينة هو 20 مشاهدة . أي

أن عدد الإشارات السالبة $k = 8$. مستوى معنوية $\alpha = 0.05$ نرفض H_0 إذا كانت :

$$P(K \leq k | n, .5) \leq \alpha .$$

نحسب :

$$P(K \leq 8 | 20, 0.5) \\ = \sum_{x=0}^8 b(x; 20, 0.5) = 0.252.$$

حيث 0.252 تستخرج من جدول ذي الحدين في ملحق (١) عند $n = 20$ و $p = 0.5$. بما أن 0.252 أكبر من 0.05 فإننا نقبل H_0 .

مثال (١٢-١٠) في مصنع للسكر تستخدم إحدى الماكينات ل تعبئة السكر في أكياس بحيث أن وزن كل كيس 5 كيلو . اختيرت عينة عشوائية من الحجم $n = 10$ من هذه الماكينة والأوزان معطاة في جدول (١٢-٢٤) .

المطلوب اختبار فرض العدم $H_0 : M = 5$ ضد الفرض البديل $H_0 : M_0 < 5$ وذلك عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢-٢٤)

4.1	4.2	4.3	4.5	4.7	4.8	4.9	4.9	5.6	4.9
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

الحل . لحساب عدد الإشارات الموجبة نحدد إشارات الفروق بين مشاهدات العينة والوسط حيث وجدت كالآتي :

-----+-----

أي أن عدد الإشارات الموجبة $k = 1$. لمستوى معنوية $\alpha = 0.05$ نرفض H_0 إذا كانت :

$$P(K \leq k | n, 0.5) \leq \alpha .$$

نحسب :

$$P(K \leq 1 | 10, 0.5) \\ = \sum_{x=0}^1 b(x; 10, 0.5) = 0.011.$$

حيث 0.011 تستخرج من جدول ذي الحدين في ملحق (١) عند $n = 10$, $p = .5$. وبما أن 0.011 أقل من 0.05 نرفض H_0 .

للعينات من الحجم 12 أو أكبر فإنه يمكننا استخدام التوزيع الطبيعي كتقريب لتوزيع ذي الحدين حيث :

$$z = \frac{(k \pm .5) - 0.5 n}{0.5 \sqrt{n}}$$

هي قيمة لمغير عشوائي Z تقريباً يتبع التوزيع الطبيعي القياسي. سوف يستخدم $k + 0.5$ إذا كانت k أقل من $\frac{n}{2}$ و $k - 0.5$ إذا كانت k أكبر من $\frac{n}{2}$. في مثال (٩-١٢) فإن :

$$z = \frac{(k \pm 0.5) - 0.5n}{0.5\sqrt{n}}$$

$$= \frac{(8 + 0.5) - 0.5(20)}{0.5\sqrt{20}} \approx -0.67.$$

لمستوى معنوية $\alpha = 0.05$ فإن $z_{0.05} = 1.645$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣). منطقة الرفض $z < -1.645$. وبما أن z المحسوبة (-0.67) تقع في منطقة القبول لقبول H_0 .

(٧-٢) اختبار الإشارة لعنتين مرتبطتين (عينة مزدوجة)

The Sign-Test for Two related Sample

يمكن تعديل اختبار الإشارة لعينة واحدة واستخدامه في حالة العنتين المرتبطتين عندهما لا يتحقق الاعتدال. بفرض أن لدينا n من أزواج المشاهدات المستقلة $(x_i, y_i); i = 1, 2, \dots, n$ فإننا نستبدل كل زوج من المشاهدات بإشارة موجبة إذا كانت الملاحظة الأولى أكبر من الملاحظة الثانية وإشارة سالبة إذا كانت الملاحظة الأولى أصغر من الملاحظة الثانية ونعمل الزوج الذي فيه المشاهدات متساويتان. لاختبار فرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D \neq 0$ أو الفرض البديل $H_1: \mu_D < 0$ أو الفرض البديل $H_1: \mu_D > 0$ نتبع الخطوات المبينة في البند السابق.

مثال (٩-١٢) يعطي الجدول (٢٥-١٢) كمية مركب في دم عينة عشوائية من 21 حيوان قبل وبعد إعطائهم دواء لتدعيم النقص في دورة ما. المطلوب اختبار فرض العدم $H_0: \mu_D = 0$ ضد الفرض البديل $H_1: \mu_D < 0$ وذلك عند مستوى معنوية $\alpha = 0.05$.

جدول (٢٥-١٢)

الحيوان	قبل x_i	بعد y_i	إشارة $(x_i - y_i)$
1	2.5	2.6	-
2	1.2	1.5	-
3	2.9	2.9	0
4	3.1	2.0	+
5	3.1	2.3	+
6	1.1	1.5	-
7	1.5	1.6	-

8	4.1	3.1	+
9	2.1	1.4	+
10	2.4	2.5	-
11	1.3	1.4	-
12	2.8	2.9	-
13	3.5	2.4	+
14	3.6	2.1	+
15	1.1	1.3	-
16	1.6	1.7	-
17	4.2	3.2	+
18	2.2	1.5	+
19	2.5	2.1	+
20	1.3	1.1	+
21	1.3	1.5	-

الحل . إشارات الفروق بين أزواج المشاهدات (x_i, y_i) ، $i = 1, 2, \dots, n$ معطاة في جدول (١٢-٢٥) وبعد إهمال الفرق المساوي للصفر وتعديل حجم العينة إلى $n = 20$ فإن عدد الإشارات الموجبة $k = 10$. نرفض H_0 عند مستوى معنوية $\alpha = 0.05$ إذا كانت :

$$P(K \leq k | n, 0.5) \leq \alpha .$$

نحسب : $P(K \leq 10 | 20, 0.5)$

$$= \sum_{x=0}^{10} b(x; 20, 0.5) = 0.588 .$$

حيث أن القيمة 0.588 تستخرج من جدول ذي الحدين في ملحق (١) عند $n = 20$, $p = 0.5$.
وعما أن 0.588 أكبر من 0.05 فإننا نقبل H_0 .

مثال (١٢-١٢) : ترغب شركة في اختبار مدي اهتداء نوعين من إطارات السيارات A , B . ولهذا قامت بتركيب الإطارات من النوع B وبعد استخدامها نفس المسافة وقياس مدي الاهتداء حصلت الشركة على البيانات المعطاة في جدول (١٢-٢٦) والمطلوب اختبار فرض العدم $H_0 : \mu_D = 0$ ضد الفرض البديل $H_1 : \mu_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

السيارة	1	2	3	4	5	6	7	8	9	10
A x_i	23	20	26	25	48	26	25	24	16	20
B y_i	20	30	16	33	23	24	8	21	13	18
إشارة $x_i - y_i$	+	-	+	-	+	+	+	+	+	+

الحل . من جدول (١٢-٢٦) فإن عدد الإشارات السالبة $k=2$. مستوى معنوية $\alpha = 0.05$ نرفض H_0 إذا كان :

$$P(K \leq k | n, 0.5) \leq \frac{\alpha}{2} .$$

نحسب :

$$\begin{aligned} P(K \leq 2 | 10, 0.5) \\ = \sum_{x=0}^2 b(x; 10, 0.5) = 0.055. \end{aligned}$$

حيث 0.055 مستخرجة من جدول ذي الخلفين في ملحق (١) عند $n = 10$, $p = 0.5$ وبما أن 0.055 أكبر من 0.025 . فإننا نقبل H_0 .

(١٢-٨) اختبار إشارة الرتب The Signed - ranks Test

يعتمد اختبار الإشارة لعينة واحدة والذي تناولناه في البند (١٢-٦) على الفرق بين قيم مشاهدات العينة والوسيط المفترض مع إغفال قيمة الفروق والذي يؤدي إلى أضعاف الاختبار . لذلك أفرح العالم Wilcoxon اختباراً لاملعياً آخر أطلق عليه اسمه يعتمد على إشارة الفرق وقيمة الفرق حيث يعطى وزناً أكبر للإشارة التي تصاحب فرقاً كبيراً والعكس صحيح . يشترك هذا الاختبار مع اختبار الإشارة في أنه يمكن أن يستخدم كاختبار للمتوسط عندما يكون المجتمع موضع الدراسة متماثل .

تكون البيانات اللازمة للتحليل من عينة عشوائية من الحجم n من المشاهدات المستقلة x_1, x_2, \dots, x_n والمختارة من مجتمع متصل. فرض العدم والفرض البديل سوف يكونان على الشكل :

$$H_0 : M = M_0,$$

$$H_1 : M \neq M_0.$$

حساب قيمة الإحصاء الذي يعتمد عليه قرارنا ، عند مستوى معنوية α ، تتبع الخطوات التالية :

$$(أ) \quad D_i = (x_i - M_0); i = 1, 2, \dots, n \quad \text{نحدد إشارة وقيمة الفروق}$$

(ب) إذا كان الفرق مساوياً للصفر تستبعد الملاحظة من التحليل ويعدل حجم العينة بطرح عدد يساوى عدد المشاهدات التي تساوي الوسيط .

(ج) نعمل إشارة الفروق مؤقتا ونرتب الفروق تصاعديا بمعنى آخر نعطى رتب للقيم $|D_i|$ أي القيم المطلقة للفروق وإذا كان هناك فروق متساوية ، أي تداخلات ties ، فإننا نعطيهما متوسط الرتب التي كانت ستأخذها لو أنها كانت مختلفة .

(د) تعاد الإشارات إلى الرتب المناظرة ونوجد مجموع رتب الإشارات السالبة ونرمز له بالرمز t_- ونوجد مجموع رتب الإشارات الموجبة ونرمز له بالرمز t_+ ويمكن إيجاد كل منهما بدلالة الآخر من العلاقة :

$$t_+ = \frac{n(n+1)}{2} - t_-$$

نحسب مجموع الرتب السالبة أو مجموع الرتب الموجبة أيهما أقل ونرمز له بالرمز t^* أي أن:

$$t^* = \min (t_+, t_-).$$

والتي تمثل قيمة لإحصاء . يستخدم الجدول في ملحق (١٩) لاستخراج القيم الحرجة لهذا الإحصاء لعينات من الحجم 3 وحتى الحجم 25 وذلك لاختبار من جانب أو جانبيين . سنرمز للقيم الجدولية بالرمز $d(n, \alpha)$, $d(n, \alpha')$ عندما يكون الاختبار من جانب أو جانبيين على التوالي. للبدليل $H_0 : M \neq M_0$ و مستوى معنوية α نرفض H_0 إذا كان $t^* < d(n, \alpha)$. للفرض البدليل $H_0 : M > M_0$ اهتمامنا سوف يكون في مجموع الرتب السالبة t_- . مستوى معنوية α نرفض H_0 إذا كان $t_- < d(n, \alpha')$. للفرض البدليل $H_0 : M < M_0$ اهتمامنا سوف يكون في مجموع الرتب الموجبة t_+ . مستوى معنوية α نرفض H_0 إذا كان $t_+ < d(n, \alpha')$.

مثال (١٢-١٣) اختبرت عينة عشوائية من 15 شخصا عن يمتلكون مولا في مدينة ما . وقد تم سؤالهم عن مقدار الزيادة في فاتورة الضرائب السنوية بالدولار. البيانات التي تم الحصول عليها معطاة في جدول (١٢-٢٧) المطلوب اختبار فرض العدم $H_1 : M = 500$ ضد الفرض البدليل $H_1 : M \neq 500$ عند مستوى معنوية $\alpha = 0.05$.

جدول (١٢ - ٢٧)

مقدار الزيادة	$D_i = (x_i - 500)$	رتب $ D_i $	مضروبا في إشارة D_i
494.4	-5.6	8	-8
510.8	+10.8	13	+13
487.5	-12.5	14	-14
493.2	-6.8	11	-11
502.6	+2.6	4	+4
500.0	0	يستهج	-
495.9	-4.1	6	-6
498.2	-1.8	3	-3
501.6	+1.6	2	+2
497.3	-2.7	5	-5
492.0	-8.0	12	-12

504.3	+4.3	7	+7
499.2	-0.8	1	-1
493.5	-6.5	10	-10
505.8	+5.8	9	+9

الحل . من جدول (١٢-٢٧) ومن العمود الأخير (على اليمين) نحسب مجموع الرتب السالبة والموجبة كالتالي :

$$t_+ = 8 + 14 + 11 + 6 + 3 + 5 + 12 + 1 + 10 = 70,$$

$$t_- = 13 + 4 + 2 + 7 + 9 = 35$$

وحيث أن هناك فرق مساوياً للصفر فإننا نعمله ونعدل حجم العينة تبعاً لذلك أي أن $n=14$.

$$t^* = \min (t_+ , t_-) = 35.$$

لمستوى معنوية $\alpha = 0.05$ فإن $d(14, 0.049) = 22$ حيث $\alpha^* = 0.049$ وعند $n=14$ من

الجدول في ملحق (١١) وبما أن $t^* = 35$ أكبر من 22 نقبل H_0 .

عندما تكون n أكبر من 25 ، فإننا لا نستطيع استخدام الجدول في ملحق (١١) لتقدير المعنوية للقيمة المحسوبة للإحصاء.

للعينات الكبيرة فإن :

$$t^* = \frac{t^* - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}.$$

تمثل قيمة للإحصاء T الذي تقريبا يتبع التوزيع الطبيعي القياسي . للاختبارات من جنس واحد فإنه يمكن استبدال t^* في صيغة t^* بالقيمة t_+ أو t_- حسب الفرض المستخدم . في حالة وجود تداعلات فلا بد من تصحيح قيمة الإحصاء t^* . بفرض أن u تمثل عدد الفروق المطلقة المتساوية لرتبة لا تساوي الصفر فإن معامل التصحيح سوف يكون كالتالي :

$$\frac{\sum u^3 - \sum u}{48}.$$

حيث يطرح هذا المقدار من المقام تحت الجذر من صيغة t^* . وعلى ذلك يصبح المقام في صيغة t^* تحت الجذر هو :

$$\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{\sum u^3 - \sum u}{48}}.$$

سوف نوضح كيفية تصحيح التداخلات باستخدام البيانات في الجدول (١٢ - ٢٨) .

جدول (١٢-٢٨)

$ D_i $	2	2	5	6	6	6	7	7	8	8	8	8
الرتب	1.5	1.5	3	5	5	5	7.5	7.5	10.5	10.5	10.5	10.5

من جدول (١٢ - ٢٨) يمكن الحصول على الجدول (١٢ - ٢٩) حيث u هو عدد القيم في كل فئة لها تداخلات .

جدول (١٢-٢٩)

U	2	3	2	4
U^3	8	27	8	256

$$\Sigma u^3 = 299 \quad , \quad \Sigma u = 11$$

وعلى ذلك معامل التصحيح للتداخلات هو :

$$\frac{\Sigma u^3 - \Sigma u}{48} = \frac{299 - 11}{48} = 6.$$

(١٢ - ٩) اختبار Mann - Whitney - Wilcoxon

يشترط في اختبار u الذي يخلص الفرق بين متوسطي مجتمعين ، الذي تناولناه في الفصل التاسع ، أن المجتمعين اللذين اخترنا منهما العيّتين يتبعان توزيعاً طبيعياً . عندما لا يتوفر هذا الشرط فإن اختبار Mann-Whitney يكون البديل .

تكون البيانات اللازمة للتحليل من عينة عشوائية من الحجم n_1 من المشاهدات x_1, x_2, \dots, x_{n_1} من المجتمع الأول المتصل . أيضاً نختار عينة عشوائية أخرى من الحجم n_2 من المشاهدات y_1, y_2, \dots, y_{n_2} من المجتمع الثاني المتصل ويشترط أن تكون العيّتين مستقلتين . فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : المجتمعين لهما نفس التوزيع .

H_1 : قيم x 's تنجه لأن تكون أصغر من قيم y 's .

لإجراء الاختبار نقوم بدمج مشاهدات العيّتين معا في عينة واحدة ثم نقوم بترتيب المشاهدات تصاعدياً فنعطي الرتبة 1 لأصغر مشاهدة والرتبة 2 للمشاهدة التي تليها في العينة وهكذا حتى المشاهدة الأخيرة والتي تمثل أكبر قيمة حيث تعطى الرتبة $n_1 + n_2$. إذا ظهرت مشاهدات متساوية في العينة (تداخلات) فإننا نرتب المشاهدات كما لو أنها ليست فيها مشاهدات متساوية في العينة ثم نحسب الوسط الحسابي لرتب المشاهدات في فئة المشاهدات المتساوية في القيمة ونعتبر الوسط الحسابي رتبة لكل مشاهدة في الفئة . نحسب القيمة :

$$w = s - \frac{n_1(n_1 + 1)}{2}$$

حيث s تمثل مجموع الرتب للعينة المختارة من المجتمع الأول و w قيمة للإحصاء W وذلك تحت فرض أن H_0 صحيح . مستوى معنوية α فإن منطقة الرفض $W < w_\alpha$ حيث w_α هي القيمة الحرجة للإحصاء W وتستخرج من الجدول في ملحق (١٢) عند n_1, n_2 ومستويات معنوية مختلفة. للفرض البديل H_0 : قيم x 's تتجه لأن تكون أكبر من قيم y 's فإن منطقة الرفض $W > w_{1-\alpha}$ تحسب كالتالي :

$$w_{1-\alpha} = n_1 n_2 - w_\alpha$$

للفرض البديل H_1 : المجتمعان يختلفان بالنسبة للموقع ، فإن منطقة الرفض $W > w_{1-\frac{\alpha}{2}}$ أو

$$W < w_{\frac{\alpha}{2}} \text{ حيث :}$$

$$w_{1-\frac{\alpha}{2}} = n_1 n_2 - w_{\frac{\alpha}{2}}$$

مثال (١٢-١٤) يعطى الجدول (٣٠-١٢) أزمنة الفشل لنوعين من الأجهزة الإلكترونية والمطلوب اختبار فرض العدم H_0 : المجتمعين لهما نفس التوزيع ضد الفرض البديل H_1 : المجتمعين ليس لهما نفس التوزيع (عند مستوى معنوية $\alpha = 0.05$).

جدول (٣٠-١٢)

x	23	261	87	7	120	14	62	47	225	71	246	21
y	55	320	56	104	220	239	47	246	176	182	33	

من الجدول (٣١-١٢) يتم حساب مجموع الرتب للعينة الأولى وهي $s=124$.

جدول (٣١-١٢)

	x	x	x	x	y	x	y	y	x	x	x	
	7	14	21	23	33	47	47	55	56	62	71	87
الرتب	1	2	3	4	5	6.5	6.5	8	9	10	11	12
	y	x	y	y	y	x	y	x	y	x	y	
	104	120	176	182	220	225	239	246	246	261	320	
الرتب	13	14	15	16	17	18	19	20.5	20.5	22	23	

وعلى ذلك فإن :

$$\begin{aligned} w &= s - \frac{n_1(n_1 + 1)}{2} \\ &= 124 - \frac{12(12 + 1)}{2} \\ &= 124 - 78 = 46. \end{aligned}$$

من الجدول في ملحق (١٢) فإن $w_{0.025}=34$ عند $n_1=12$ و $n_2=11$ وبما أن :

$$w_{1-\frac{\alpha}{2}} = n_1 n_2 - w_{\frac{\alpha}{2}}$$

فإن :

$$w_{0.975} = (12)(11) - 34 = 98.$$

منطقة الرفض $W > 98$ أو $W < 34$ وبما أن $w = 46$ تقع في منطقة القبول نقبل H_0 .
عندما n_1, n_2 أكبر من 20 فإننا لا نستطيع استخدام الجدول في ملحق (١٢) وفي هذه الحالة يمكن حساب القيمة :

$$z = \frac{w - \frac{n_1 n_2}{2}}{\sqrt{(n_1 n_2)(n_1 + n_2 + 1)/12}}$$

والتي تمثل قيمة للمتغير العشوائي Z وهو تقريباً يتبع التوزيع الطبيعي القياسي وذلك تحت فرض أن H_0 صحيح .

بالنسبة لمشكلة التداخلات والتي قد تحدث داخل كل مجموعة أو بين قيم المجموعتين فقد تم إثبات أن التداخلات داخل المجموعة ليس لها تأثير على قيمة الإحصاء ولكن وجود تداخلات بين المجموعتين يؤثر على النتائج . في هذه الحالة لا بد من عمل تصحيح لقيمة z . بفرض أن u ترمز لعدد التداخلات لرُب معطاة فإن معامل التصحيح للتداخلات يحسب من الصيغة التالية :

$$\frac{n_1 n_2 (\sum u^3 - \sum u)}{12(n_1 + n_2)(n_1 + n_2 - 1)}$$

والتي تطرح من المقام في صيغة z تحت الجذر . وعلى ذلك المقام في صيغة z يصبح :

$$\sqrt{\frac{(n_1 n_2)(n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 (\sum u^3 - \sum u)}{12(n_1 + n_2)(n_1 + n_2 - 1)}}$$

Kruskal - Wallis (١٢-١٠) اختبار

يعتبر اختبار $Kruskal - Wallis$ من أكثر الاختبارات اللامعلمية شوعاً لإجراء تحليل التباين في حالة التصنيف الأحادي. تتكون البيانات اللازمة للتحليل من k من العينات العشوائية من الحجم n_1, n_2, \dots, n_k على أن تكون المشاهدات مستقلة سواء بين أو داخل المعالجات كما أن المجتمعات التي اختيرت منها العينات تكون من النوع المتصل . فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : التوزيعات للمجموعات التي عددها k متطابقة .

H_1 : التوزيعات للمجموعات التي عددها k ليس لها نفس الوسيط .

في هذا الاختبار نقوم بدمج مشاهدات العينات في عينة واحدة وإعطاء رتب لهذه المشاهدات حيث R_i ترمز إلى مجموع الرتب للمشاهدات التي تنتمي إلى المجموعة رقم i والتي عدد مشاهداتها n_i حيث أن $N = \sum_{i=1}^k n_i$ تمثل العدد الكلي للمشاهدات الناتجة من k من

العينات . قيمة الإحصاء الذي يعتمد عليه قرارنا هو :

$$h = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{1}{n_i} \left[R_i - \frac{n_i(N+1)}{2} \right]^2$$

والتي يمكن تبسيطه بالصيغة التالية :

$$h = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1).$$

حيث أن h هو قيمة للإحصاء H بالفترض أن H_0 صحيح . تستخرج القيم الحرجة للإحصاء H من الجدول في ملحق (١٣) عند مستويات مختلفة من المعنوية بشرط أن $n_i \leq 5, i = 1, 2, 3$. إذا كان عدد العينات أو عدد المشاهدات داخل كل عينة غير متوفر في الجدول فقد وجد بالبرهان أن H تقريبا تتبع توزيع χ^2 تحت شرط أن عدد المشاهدات في كل عينة لا تقل عن 5 أي أن $n_i \geq 5, i = 1, 2, \dots, k$. مستوى معنوية α فإن منطقة الرفض $\chi_\alpha^2 > \chi_\alpha^2$ حيث أن χ_α^2 تستخرج من جدول توزيع χ^2 عند مستوى معنوية α ودرجات حرية $k-1$. نرفض H_0 إذا وقعت h المحسوبة في منطقة الرفض .

في حالة وجود تداعلات لا بد من تصحيح الإحصاء H إلى الإحصاء H' حيث :

$$H' = \frac{H}{C},$$

$$C = 1 - \frac{\sum (u_i^3 - u_i)}{N^3 - N}$$

حيث u_i هو عدد القيم في كل فئة بما قيم متساوية و N هي القيمة الناتجة من دمج العينات

التي عددها k أي أن $N = \sum_{i=1}^k n_i$. مستوى معنوية α فإن منطقة الرفض $H' > \chi_\alpha^2$ حيث :

χ^2_{α} يستخرج من جدول توزيع χ^2 في ملحق (٥) بدرجات حرية $k-1$. إذا وقعت h' المحسوبة ($h' = \frac{h}{C}$) في منطقة الرفض نرفض H_0 . إذا كانت القيم المتداخلة قليلة في العينات فإن H' تقترب من H .

مثال (١٢-١٥) استخدمت ثلاثة طرق تعليمية مختلفة لتعليم ثلاثة مجموعات متشابهة من الطلبة وكانت درجات الامتحان النهائي معطاة في جدول (١٢-٣٢)

جدول (١٢-٣٢)

الجموعة A	20	37	39	41	45
الجموعة B	43	46	48	53	
الجموعة C	31	38	44		

هل توجد فروق معنوية بين الطرق الثلاثة؟ وكذلك عند مستوى معنوية $\alpha = 0.05$.
الحل .

H_0 : توزيعات المجتمعات الثلاثة التي اختبرت فيها العينات الثلاثة متطابقة .

H_1 : توزيعات المجتمعات الثلاثة ليس لها نفس الوسيط.

للحصول على القيمة h تم دمج الثلاث عينات معا ووضع رتب للعينات المشتركة وتحديد رتب كل عينة والنتائج معطاة في الجدول (١٢-٣٣).

جدول (١٢-٣٣)

	الترتيب					R_i
A	1	3	5	6	9	24
B	7	10	11	12		40
C	2	4	8			14

وعلى ذلك فإن قيمة h هي :

$$\begin{aligned}
 h &= \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i}{n_i} - 3(N+1) \\
 &= \frac{12}{N(N+1)} \left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} \right] - 3(N+1) \\
 &= \frac{1}{13} \left[\frac{24^2}{5} + \frac{40^2}{4} + \frac{14^2}{3} \right] - 3(13)
 \end{aligned}$$

$$= \left(\frac{1}{13} \cdot 580.53 \right) - 39$$

$$= 44.66 - 39 = 5.66.$$

لمستوى معنوية $\alpha = 0.05$ القيمة الحرجة للإحصاء H والمستخرجة من الجدول في ملحق (

١٣) هي 5.6308 عند $n_1 = 5, n_2 = 4, n_3 = 3$. منطقة الرفض $H > 5.6308$.

وبما أن $h = 5.66$ تقع في منطقة الرفض نرفض H_0 .

مثال (١٦-١٦) في دراسة على نوع معين من الطحالب في بحيرة صغيرة قام باحث بمقارنة

النسبة المئوية لعدد الطحالب من نوع ما في منطقة محددة خلال فصل الربيع ومتصف الصيف

وآخر الخريف والبيانات معطاة في جدول (١٦-٣٤).

جدول (١٦-٣٤)

الربيع %	9.6	11.2	11.6	11.7	12.8	12.9	15.8	22.7	24.6	32.5
متصف الصيف %	4.8	7.6	7.6	9.2	9.6	21.1	24.6	25.6	26.4	32.8
آخر الخريف %	5.4	6.5	7.1	8.0	8.8	9.5	10.2	10.7	11.3	11.7

النتائج اللازمة لحساب h في جدول (١٦-٣٥)

جدول (١٦-٣٥)

											R _i
الربيع	11.5	15	17	18.5	20	21	22	24	25.5	29	203.5
منتصف الصيف	1	5.5	5.5	9	11.5	23	25.5	27	28	30	166
آحر الخريف	2	3	4	7	8	10	13	14	16	18.5	95.5

$$h = \frac{12}{N(N+1)} \left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} \right] - 3(N+1)$$

$$= \frac{12}{30(31)} \left[\frac{203.5^2}{10} + \frac{166^2}{10} + \frac{95.5^2}{10} \right] - 3(31)$$

$$= 7.759.$$

ولوجود تداخلات لابد من تحويل h إلى h' . نحسب $u^3 - u$ لكل تداخل ثم نحسب $\Sigma(u^3 - u)$ كالتالي :

$$(8-2) + (8-2) + (8-2) + (8-2) = 24.$$

معامل التصحيح بحسب من الصيغة التالية :

$$1 - \frac{\Sigma(u^3 - u)}{N^3 - N}$$

$$= 1 - \frac{24}{27000 - 30} = 0.9991.$$

وعلى ذلك :

$$h' = \frac{7.759}{0.9991} = 7.76599.$$

$\chi^2_{0.05} = 5.992$ والمستخرجة من جدول توزيع χ^2 في ملحق (٥) عند درجات حرية $k-1=3-1=2$. منطقة الرفض $X^2 > 5.992$. وبما أن $h' = 7.76599$ تقع في منطقة الرفض نرفض H_0 .

Test Runs (١١-١٢) اختبار الدورات

يفيد هذا الاختبار في مجالات كثيرة منها المشاكل البيولوجية . فقد يرغب باحث ما في معرفة ما إذا كان عدد حالات الإصابة بمرض الملاريا ، مثلاً ، يتغير عشوائياً من سنة إلى سنة أخرى أم أن هناك عوامل غير عشوائية تؤدي إلى نقص أو زيادة عدد حالات الإصابة بهذا المرض . لإجراء الاختبار نفترض أن لدينا متابعة من المشاهدات المسجلة تبعاً لترتيب حدوثها وأن المشاهدات يمكن تقسيمها إلى نوعين (ليكن a, b) . يعتمد هذا الاختبار على متغير يطلق عليه أسم الدورة run حيث تعرف بأنها مجموعة الأحداث المشاهدة التي يسبقها أو يتبعها نوع آخر مخالفها من الأحداث أو لا يتبعها أو لا يسبقها أية أحداث . عدد الأحداث في الدورة يطلق عليها طول الدورة (يمكن أن تحصى الدورة على حدث واحد) . بفرض أن لدينا البيانات التالية والتي تم توليدها على الحاسب الآلي .

0.1, 0.4, 0.2, 0.8, 0.6, 0.9, 0.3, 0.4, 0.01, 0.2

بفرض أن a تمثل الرقم الذي أقل من 0.5 و b تمثل الرقم الذي أكبر من 0.5 والتي تعطى المتابعة :

a a a b b b a a a a

في هذه الحالة لدينا ثلاث دورات . سوف نرمز لعدد الدورات بالرمز r' ، أي أن $r' = 3$.
أيضا n_1 سوف نرمز لعدد قيم a و n_2 سوف نرمز لعدد قيم b . فرض العدم والفرض البديل
سوف يكونان على الشكل

H_0 : النوعان يقعان بعشوائية .

H_1 : النوعان لا يقعان بعشوائية .

بفرض أن H_0 صحيح فإن r' هي قيمة للإحصاء R' . القيم الحرجة السفلي للإحصاء R'
تستخرج من الجدول في ملحق (١٤) والقيم الحرجة العليا للإحصاء R' تستخرج من الجدول
في ملحق (١٥) وذلك عند مستوى معنوية $\alpha = 0.05$ ، n_1, n_2 . تكون القيمة السفلي و
 r_2 القيمة العليا عند $\alpha = 0.05$ ، n_1, n_2 ، منطقة الرفض سوف تكون $R' > r_2$ أو $R' < r_1$
إذا وقعت القيمة r في منطقة الرفض نرفض H_0 . للفرض البديل H_1 : النوعان لا يقعان
بعشوائية لوجود عدد كبير من الدورات والمستوى معنوية $\frac{\alpha}{2}$ فإن منطقة الرفض $R' > r_2$.
للفرض البديل H_1 : النوعان لا يقعان بعشوائية لوجود عدد قليل من الدورات والمستوى معنوية
 $\frac{\alpha}{2}$ فإن منطقة الرفض $R' < r_1$. إذا كانت n_1, n_2 أكبر من 20 (كليهما أو أيهما) فإن
الجدولان في ملحق (١٤) و (١٥) لا يصلحان للاستخدام ولقد وجد بالبرهان أن :

$$Z = \frac{r' - \left[\frac{2n_1n_2}{n_1 + n_2} + 1 \right]}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}}$$

قيمة لتطير عشوائي Z تقريباً يتبع التوزيع الطبيعي القياسي .

مثال (١٢ - ١٧) إذا كان عدد حالات الملاريا المسجلة في منطقة ما خلال عدة سنوات
مسجلة في الجدول (١٢ - ٣٦) .

جدول (١٢ - ٣٦)

51	82	64	32	11	12	54	71	90	101	84	72	45	20	74	15
----	----	----	----	----	----	----	----	----	-----	----	----	----	----	----	----

هل عدد حالات الإصابة تأخذ شكل عشوائي أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل . H_0 : البيانات عشوائية .

H_1 : البيانات غير عشوائية .

عادة يعتبر الوسيط هو القيمة التي تستخدم لتقسيم البيانات في العينة إلى نوعين a و b . للبيانات في جدول (١٢-٣٦) فإن الوسيط 59 . سوف نرمز للقيمة التي أصغر من قيمة الوسيط بالرمز a والقيمة التي أكبر من قيمة الوسيط بالرمز b وبذلك نحصل على المتابعة التالية :

a bb aaaa bbbbbb aa b a

وعلى ذلك :

$$r' = 7, \quad n_2 = 8, \quad n_1 = 8$$

لستوى معنوية $\alpha = 0.05$ فإن $r_1 = 4$ والمستخرجة من الجدول في ملحق (١٤) عند $n_1 = 8$ و $n_2 = 8$ ، $r_2 = 14$. والمستخرجة من الجدول في ملحق (١٥) عند $n_1 = 8$ و $n_2 = 8$. منطقة الرفض $R' > 14$ أو $R' < 4$. وبما أن $r' = 7$ فإنها تقع في منطقة القبول وبذلك نقبل فرض العدم .

مثال (١٢ - ١٨) إذا كان لدينا 34 شخصا وبفرض أن M ترمز للذكر و F ترمز للإناث وكانت النتائج كالآتي :

FF MMMMMMMMMM FF M FF MMMMMMMMMM FFFFF
MMM F

المطلوب تحديد هل العينة عشوائية أم لا ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل . H_0 : العينة عشوائية.

H_1 : العينة غير عشوائية.

من بيانات العينة نجد أن :

$r' = 9$, $n_2 = 22$ (عدد M) و $n_1 = 12$ (عدد F) وبما أن $n > 20$ فإننا نستخدم

التقريب الطبيعي وعلى ذلك :

$$z = \frac{r' - \left[\frac{2n_1n_2}{n_1 + n_2} + 1 \right]}{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

$$= \frac{9 - \left[\frac{2(12)(22)}{12 + 22} + 1 \right]}{\frac{2(12)(22)[2(12)(22) - 12 - 22]}{(12 + 22)^2(12 + 22 - 1)}}$$

$$= \frac{9-16.529}{6.8373702} \\ = -1.101.$$

لمستوى معنوية $\alpha = 0.05$ و $z_{0.025} = 1.96$ والمستخرجة من جدول التوزيع الطبيعي القياسي في ملحق (٣). منطقة الرفض $Z > 1.96$ أو $Z < -1.96$ وبما أن $z = -1.101$ تقع في منطقة القبول نقبل H_0 .

(١٢-١٤) معامل ارتباط سيرمان للرتب

The Spearman Rank Correlation Coefficient

تناولنا في البند (١٠-٤) من الفصل العاشر اختبارات الفروض التي تخص معامل ارتباط المجتمع ρ تحت فرض أن X, Y متغيرين عشوائيين لهما توزيع طبيعي ثنائي. في حالة عدم تحقق الشرط السابق فإنه يمكننا استخدام معامل سيرمان كإحصاء لاختبار عدم وجود علاقة (ارتباط) بين المتغيرين X, Y . أيضا يمكننا استخدام معامل سيرمان كمقياس وصفي لقوة الارتباط بين متغيرين X, Y عندما تكون البيانات في العينة غير متوفرة في شكل بيانات رقمية ولكن يمكن تعيين رتب لها. لإجراء الاختبار نتبع الآتي:

(أ) نختار عينة عشوائية من الحجم n من أزواج المشاهدات الرقمية أو الوصفية. كل زوج من المشاهدات يمثل قراءتين مأخوذتين على نفس المفردة والمسماة وحدة الاقتران **unit of association**. أيضا قد تمثل البيانات مشاهدات مأخوذة من مجتمع ثنائي. سوف نرمز لأزواج المشاهدات كالتالي $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

(ب) نرتب قيم المشاهدات في العينة والتابعة للمتغير X تصاعدياً ونعطي رتبة لكل قيمة مشاهدة بالنسبة لكل قيم المشاهدات الأخرى. سوف نرمز لرتبة المشاهدات رقم i ، x_i ، بالرمز $r(x_i)$. عندما $r(x_i) = 1$ فهذا يعني أن x_i تمثل أقل قيمة مشاهدة من قيم المتغير X في العينة.

(ج) نرتب قيم المشاهدات في العينة والتابعة للمتغير Y تصاعدياً ونعطي رتبة لكل قيمة مشاهدة بالنسبة لكل قيم المشاهدات الأخرى. سوف نرمز لرتبة المشاهدات رقم j ، y_j ، بالرمز $r(y_j)$. عندما $r(y_j) = 1$ فهذا يعني أن y_j تمثل أقل قيمة مشاهدة من قيم المتغير Y في العينة.

(ح) عند حدوث تداخلات نعطي متوسط الرتب المتتالية بدلاً من الرتبة كالتعداد.

(خ) إذا كانت البيانات وصفية بإمكاننا تحويلها إلى رتب.

قيمة الإحصاء الذي يعتمد عليه قرارنا هو معامل ارتباط سيرمان والذي يحسب من الصيغة التالية :

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

حيث:

$$\sum d_i^2 = \sum [r(x_i) - r(y_i)]^2.$$

لكل زوج من المشاهدات وعندما تكون رتبة x نفس رتبة y (ارتباط تام طردي) ، فإن كل الفروق d_i سوف تساوي صفر وعلى ذلك $r_s = +1$. إذا كانت رتبة كل متغير داخل كل زوج من المشاهدات عكس الآخر (ارتباط تام عكسي) ، أي إذا كان :

$[r(x)=1, r(y)=n] , [r(x)=2, r(y)=n-1] , \dots , [r(x)=n, r(y)=1]$
وذلك لأزواج المشاهدات التي عددها n فإن $r_s = -1$ على سبيل المثال إذا كان لدينا أزواج المشاهدات التالية :

$$(x_i, y_i) : (12, 5), (11, 6), (10, 7), (9, 8)$$

فإن الرتب تصبح :

$$\begin{array}{cccc} r(x_i) : & 4 & 3 & 2 & 1 \\ r(y_i) : & 1 & 2 & 3 & 4 \end{array}$$

وعلى ذلك $\sum d_i^2$ سوف تكون :

$$(3)^2 + (1)^2 + (-1)^2 + (-3)^2 = 20.$$

وبالتعويض في معادلة سيرمان فإن :

$$\begin{aligned} r_s &= 1 - [(6) (20) / (4) (15)] \\ &= 1 - 2 = -1. \end{aligned}$$

معامل ارتباط سيرمان لا يمكن أن يزيد عن 1 ولا يمكن أن يقل عن -1 . فرض العدم والفرض البديل سوف يكونان على الشكل :

H_0 : المتغيرين مستقلين .

H_1 : توجد علاقة بين المتغيرين في نفس الاتجاه أو الاتجاه المعاكس .

بفرض أن H_0 صحيح فإن r_s تمثل قيمة للإحصاء R_s الذي له توزيع احتمالي . القيم الحرجة

$r_{s,\alpha}^*$ للإحصاء R_s تستخرج من الجدول في ملحق (١٦) لعينات من الحجم 4 وحتى الحجم 30 عن مستويات مختلفة من المعنوية . مستوى معنوية α فإن منطقة الرفض $R_s > r_{s,\alpha}^*$ أو

$R_s < -r_{s,\alpha}^*$. إذا وقعت r_s في منطقة الرفض فإننا نرفض H_0 . للفرض البديل H_1 :
توجد علاقة بين المتغيرين في نفس الاتجاه فإن منطقة الرفض $R_s > r_{s,\alpha}^*$ وذلك عند مستوى
معنوية α . للفرض البديل H_1 : توجد علاقة بين المتغيرين في اتجاه معاكس فإن منطقة الرفض
 $R_s < -r_{s,\alpha}^*$ وذلك عند مستوى معنوية α .

القرارات السابقة تستخدم عندما لا يكون هناك تداخل أو أن يكون عددها صغيراً .
عندما يكون هناك تداخل وإذا كان عددها كبيراً (العدد الصغير للتدخلات لا يؤثر على r_s)
(فيجب إجراء تصحيح على r_s ونحتاج جداول خاصة لإجراء الاختبار سوف لا تعرض هنا .
عندما يكون حجم العينة كبيراً (أكبر من 30) فإننا لا نستطيع استخدام الجداول ولكن تم إثبات
أن :

$$z = r_s / \sqrt{n-1}$$

قيمة للمتغير العشوائي Z والذي تقريباً يتبع التوزيع الطبيعي القياسي وذلك بفرض أن
 H_0 صحيح .

مثال (١٢ - ١٩) لدراسة العلاقة بين الهيموجلوبين X (مقاساً mg/100 ml) وعدد كرات
الدم الحمراء Y بالمليون لكل ملليمتر مكعب ، اختيرت عينة عشوائية من 12 ذكر بالغ من
مجمع ما وتم قياس تركيزات الهيموجلوبين وعدد كرات الدم الحمراء لكل مفردة والبيانات
معطاة في جدول (١٢ - ٣٧) .

جدول (١٢ - ٣٧)

الشخص	الهيموجلوبين		كرات الدم الحمراء		d .	d ²
	x	رتب x	y	رتب y		
1	15.2	7.5	5.1	9	-1.5	2.25
2	16.4	12	5.4	11	1	1
3	14.2	2	4.5	4	-2	4
4	13.0	1	4.2	1	0	0
5	14.5	3	4.3	2.5	0.5	0.25
6	16.1	11	6.1	12	-1	1
7	15.2	7.5	5.2	10	-2.5	6.25
8	14.8	5	4.3	2.5	2.5	6.25
9	15.7	10	4.7	6	4	16
10	14.9	6	4.8	7.5	-1.5	2.25
11	15.6	9	4.6	5	4	16
12	14.7	4	4.8	7.5	-3.5	12.25

المطلوب اختبار لفرض العدم H_0 : المتغيرين مستقلين ضد الفرض البديل H_1 : توجد علاقة بين المتغيرين في نفس الاتجاه أو الاتجاه المعاكس وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل . $\sum d_i^2 = 67.5$ وعلى ذلك فإن :

$$\begin{aligned} r_s &= 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \\ &= 1 - \frac{6(67.5)}{12(144 - 1)} \\ &= 1 - 0.2360139 = 0.763986. \end{aligned}$$

$r_s^* = 0.5804$ والمستخرجة من الجدول في ملحق (١٦) عند مستوى معنوية

$n = 12$, $\frac{\alpha}{2} = 0.025$ منطقة الرفض $R_s > 0.5804$ أو $R_s < -0.5804$. وبما أن

$r_s = 0.763986$ تقع في منطقة الرفض نرفض H_0 .

مثال (١٢ - ٢٠) يعطى الجدول (١٢ - ٣٨) تقديرات 10 طلاب في كل من الإحصاء والرياضيات .

جدول (١٢ - ٣٨)

جيد	جيد	ممتاز	جيد	جيد	جيد	مقبول	جيد	جيد	تقديرات الإحصاء
جيد	جيد	ممتاز	جيد	جيد	جيد	مقبول	جيد	جيد	تقديرات الرياضيات

أخبر لفرض العدم H_0 : المتغيرين مستقلين .

ضد الفرض البديل :

H_1 : توجد علاقة بين المتغيرين في نفس الاتجاه .

وذلك عند مستوى معنوية $\alpha = 0.05$.

الحل . من جدول (١٢ - ٣٨) يمكن الحصول على جدول (١٢ - ٣٩) .

جدول (١٢ - ٣٩)

الاجموع	4	4	9.5	4	9.5	7.5	4	1	4	7.5	رب x
	7.5	7.5	10	1.5	7.5	4	7.5	1.5	4	4	رتب y
	-3.5	-3.5	-0.5	2.5	2	3.5	-3.5	-0.5	0	3.5	D _i
	12.25	12.25	0.25	6.25	4	12.25	12.25	0.25	0	12.25	d _i ²

وعلى ذلك :

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(72)}{10(100 - 1)}$$

$$= 1 - 0.4363636 = 0.5636363.$$

$r_{s,0.05}^* = 0.5515$ والمستخرجة من الجدول في ملحق (١٦) عند مستوى معنوية $n = 10, \alpha = 0.05$. منطقة الرفض $R_s > 0.5515$. وبما أن $r_s = 0.5636363$ تقع في منطقة الرفض نرفض H_0 .

تأريـن :

١- اختير 178 رقم من أحد الجداول عشوائياً وكان التوزيع التكراري لهذه الأرقام كما في الجدول التالي :

الرقم	1	2	3	4	5	6	7	8	9
التكرار	17	18	22	20	15	29	22	20	15
المشاهد									

اختبر فرض العلم ، عند مستوى معنوية $\alpha = 0.05$ ، أن البيانات السابقة يمكن توليفها بالتوزيع المتظم .

٢- ألقيت عمله 100 مرة وتم الحصول على 63 صورة و 37 كتابة . هل هذه النتائج تتفق مع الفرض القائل أن العملة متزنة ؟ استخدم مستوى معنوية $\alpha = 0.05$.

٣- طبقا لنظرية في علم الوراثة فإن تجين الورود الحمراء والورود الصفراء ينتج ورود حمراء بنسبة 50% وورود برتقالية بنسبة 25% وورود صفراء بنسبة 25% . لاختبار صحة هذه النظرية أجريت 164 عملية تجين فتم الحصول على 29 وردة حمراء و 90 وردة برتقالية و 45 وردة صفراء . هل تدل هذه البيانات على صحة النظرية . استخدم مستوى معنوية $\alpha = 0.05$.

٤- ألقيت زهرة 120 مرة متتالية وتم الحصول على النتائج التالية :

عدد النقاط	1	2	3	4	5	6
التكرار المشاهد	18	21	17	22	23	19

هل يمكن القول بأنه هذه الزهرة غير متحيزة باستخدام مستوى معنوية $\alpha = 0.01$.

٥- أقيمت أربع قطع نقود عشرين مرة وتم الحصول على النتائج التالية :

عدد الصور	0	1	2	3	4
التكرار المشاهد	3	4	5	5	3

المطلوب استخدام اختبار مربع كاي لجودة التوفيق لاختبار ما إذا كانت النتائج تتفق مع توزيع

ذي الحدين باحتمال نجاح $\frac{1}{2}$ وذلك عند مستوى معنوية $\alpha = 0.01$.

٦- اختبرت ثلاثة كرات من إبناء يحتوي على 5 كرات حمراء و 3 كرات خضراء وتم تسجيل

العدد x الذي يمثل عدد الكرات الحمراء ثم أعيدت الكرات في الإناء وكررت التجربة 112 مرة

. النتائج التي تم الحصول عليها في الجدول التالي :

x	0	1	2	3
التكرار المشاهد	1	31	55	25

أختبر فرض العدم ، عند مستوى معنوية $\alpha = 0.01$ ، أن البيانات التي تم الحصول عليها من

التجربة يمكن توفيقها بالتوزيع الهندسي الزائدي $(x; 8, 3, 5)$ حيث $x=0,1,2,3$.

٧- إذا كانت درجات مادة الإحصاء لفصل دراسي خاص كالآتي :

الدرجة	A	B	C	D	E
التكرار المشاهد	14	18	32	20	16

أختبر فرض العدم ، عند مستوى معنوية $\alpha = 0.01$ ، أن البيانات في الجدول السابق يمكن

توفيقها بتوزيع منتظم .

٨- أقيمت عملة حتى ظهر وجه فإذا كان X يمثل عدد المحاولات حتى ظهور وجهه . بعد

تكرار التجربة 256 مرة تم الحصول على البيانات التالية :

x	1	2	3	4	5	6	7	8
التكرار المشاهد	136	60	34	12	9	1	3	1

أختبر فرض العدم ، عند مستوى معنوية $\alpha = 0.01$ ، أن البيانات في الجدول السابق يمكن

توفيقها بالتوزيع الهندسي $g(x, \frac{1}{2})$, $x=1,2,3, \dots$.

٩- قامت شركة للتأمين على السيارات بتسجيل البيانات الخاصة بعدد الحوادث x التي

تعرضت لها السيارات المؤمن عليها والبيانات في الجدول التالي :

التكرار	10	40	100	150	200	125	75	50	30	20
المشاهد										

هل يمكن القول أن عدد الحوادث التي تتعرض لها السيارات المؤمن عليها تتبع توزيع بواسون وذلك عند مستوى معنوية $\alpha = 0.05$.

- ١٠- البيانات في الجدول التالي تعطي التوزيع التكراري لمتغير X .

(أ) وفق منحني التوزيع الطبيعي لهذه البيانات.

(ب) اختبر جودة التوفيق وذلك عند مستوى معنوية $\alpha = 0.05$.

حدود الفئة	0-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16
التكرار	15	20	35	25	35	40	20	15

- ١١- البيانات التالية تعطي عدد الوفيات بسبب تعاطي جرعات زائدة من مشروب كحلي وذلك عند أعمار مختلفة :

العمر	9-15	16-22	23-29	30-36	37-43	44-50	51-57
التكرار	40	35	32	10	13	13	4

المطلوب (أ) وفق منحني التوزيع الطبيعي لهذه البيانات؟

(ب) اختبر جودة التوفيق عند مستوى معنوية $\alpha = 0.01$.

- ١٢- في معركة حربية كان عدد المناطق m_y التي تستقبل y ضربات كالتالي :

y	0	1	2	3	4	≥ 5
m_y	229	211	93	35	7	1

اختبر فرض العدم أن عدد الضربات متغير عشوائي يتبع توزيع بواسون بمعلمة μ عند مستوى

معنوية $\alpha = 0.05$ افترض أن $\bar{x} \approx 3.25$ عند مستوى معنوية $\alpha = 0.05$.

- ١٣- إذا كانت أزيمة الحياة لـ 100 غلية ضوئية flashlight ، تم ملاحظتها موزعة كالتالي:

عدد الدقائق	0-706	706-746	746-786	786-∞
التكرار	13	36	38	13

اختبر فرض العدم أن زمن الحياة X للخلية الضوئية يتبع توزيعاً طبيعياً.

- ١٤- قام مسئول مراقبة الجودة في مصنع لإنتاج وحدات معينة باختيار 200 وحدة من ناحية

عدد السطوح التالفة وتم الحصول على البيانات التالية :

عدد المطوح التالفة	0	1	2	3	4	5	6 وأكثر
التكرار المشاهد	90	62	31	13	3	1	0

استخدم توزيع بواسون لتوفيق هذه البيانات واختبر جودة التوفيق عند مستوى معنوية $\alpha = 0.05$.

١٥- أوضحت الدراسات السابقة في مصنع للملابس أن 14% من المبيعات كانت للأشخاص الذين عمرهم أقل من 16 سنة و 38% كانت للأعمار من 16 إلى 20 سنة و 26% كانت للأعمار من 21 إلى 25 سنة و 22% كانت للأشخاص أكبر من 25 سنة. الجدول التالي يعطي النتائج التي تم الحصول عليها من عينة من 200 شخص.

العمر	أقل من 16	16-20	21-25	أكبر من 25
التكرار	22	62	60	56

ما هو فرض العدم الذي نتوقع أن نتخبره ؟ ضع استنتاجك عند مستوى معنوية $\alpha = 0.05$.

١٦- إذا كان X يمثل الزمن اللازم (بالأيام) لتصليح مكون ما في طائرة . المطلوب اختبار ما إذا كان توزيع بواسون بموسط $= 3$ ملائماً نموذج مناسب لهذا المتغير وذلك باستخدام البيانات في الجدول التالي :

زمن التصليح (بالأيام)	0	1	2	3	4	5	6	≥ 7
التكرار	1	3	7	6	10	7	6	0

١٧- يعطي الجدول التالي سرعة الإشعاع radial velocities لعدد 80 من النجوم الساطعة والمطلوب اختبار ما إذا كان التوزيع الطبيعي نموذج قياسي لهذا المتغير .

حدود الفئة	(-80,-70)	(-70,-60)	(-60,-50)	(-50,-40)	(-40,-30)
التكرار	1	2	2	2	8
حدود الفئة	(-30,-20)	(-20,-10)	(-10, 0)	(0,10)	(10,20)
التكرار	24	26	11	2	1
حدود الفئة	(20,30)				
التكرار	1				

١٨- في عينة عشوائية من 1000 شخص وجد أن 452 يعطون صوته لمرشح ما. المطلوب اختبار فرض العدم أن 50% من الأصوات تذهب إلى المرشح وذلك عند مستوى معنوية $\alpha = 0.01$.

١٩- خلال 1000 ساعة طيران لخمسين طائرة تم تسجيل عدد الطائرات m_x التي حدث لها فشل في x مكون . النتائج في الجدول التالي :

x	0	1	2	3	4	≥ 5
m_x	7	10	9	7	6	11

اختبر فرض العدم أن المتغير X يتبع توزيع بواسون بمعلمة $\mu = 2$ عند مستوى معنوية $\alpha = 0.01$.

٢٠- إذا كان عدد الصاديق المفروضة من الحجم الواحد من مسحوق منظف من الأنواع A, B, C, D, E, F هي :

النوع	A	B	C	D	E	F
عدد الوحدات المفروضة	270	308	290	312	300	320

اختبر فرض العدم أن الاختلاف بين الأنواع المختلفة يرجع إلى الصدفة باستخدام مستوى معنوية $\alpha = 0.05$.

٢١- في بحث ميداني عن التفضيل خمسة أنواع من مساحيق التنظيف A, B, C, D, E تم الحصول على البيانات التالية :

النوع	A	B	C	D	E
التكرار	187	221	193	204	195

هل تعتقد أن المستهلكين غير قادرين على التمييز بين المساحيق الخمسة ؟ عند مستوى معنوية $\alpha = 0.05$.

٢٢- في مشكلة جنينة يعتقد أن اللون البني يحدث باحتمال 25. والأبيض باحتمال 0.25 والمنقط باحتمال 0.5 .

(أ) اختبر فرض العدم أن هذا النموذج صحيح عند مستوى معنوية $\alpha = 0.01$ إذا تم الحصول على البيانات التالية :

اللون	البني	الأبيض	المنقط
التكرار	5	15	20

(ب) اختبار فرض العدم أن الاحتمالات هي $\frac{1}{9}$ للبنى و $\frac{4}{9}$ للأبيض و $\frac{4}{9}$ للمتقط وذلك عند مستوى معنوية $\alpha = 0.05$.

٢٣- إذا كان احتمال أن لاعب كرة السلة يصوب نحو الهدف هو 0.3 فإذا لعب 100 محاولة فاحصل على 20 هدف اختبار فرض العدم $H_0: P = 3$ ضد الفرض البديل $H_1: P \neq 0.3$ حيث P هو احتمال التصويب نحو الهدف عند مستوى معنوية $\alpha = 0.01$.

٢٤- اختبرت عينة عشوائية من 30 لوداً في جامعة ما وتم تصنيفهم تبعاً للجنس وعدد ساعات مشاهدة التلفزيون في خلال أسبوع. البيانات التي تم الحصول عليها معطاة في الجدول التالي :

المشاهدة	الجنس	
	ذكر	أنثى
أكثر من 25 ساعة	5	9
أقل من 25 ساعة	9	7

هل يمكن القول أن هناك علاقة بين الجنس وعدد ساعات مشاهدة التلفزيون وذلك عند مستوى معنوية $\alpha = 0.05$.

٢٥- أخذت عينة عشوائية من 200 رجل متزوج وتم تصنيفهم في الجدول التالي تبعاً للتعليم وعدد الأطفال :

التعليم	عدد الأطفال		
	0-1	2-3	أكثر من 3
بسيط	14	37	32
متوسط	19	42	17
جامعي	12	17	10

اختبر فرض العدم أن عدد الأفراد في الأسرة غير مرتبط بالتعليم وذلك عند مستوى معنوية $\alpha = 0.01$.

٢٦- لدراسة العلاقة بين امتلاك سيارة وامتلاك تلفزيون في بلد ما تم إجراء استطلاع للرأي على عينة من 20,000 شخص وقد تم الحصول على البيانات التالية :

عدد التلفزيونات التي يمتلكها الشخص	عدد السيارات التي يمتلكها الشخص			
	لا يوجد	1	2	3

لا يوجد	1000	900	100	2000
واحد	1500	2600	500	4600
أثنين أو أكثر	500	2500	400	3400

أختبر فرض العدم أن عدد السيارات التي يمتلكها الشخص مستقلة عن عدد التلفزيونات التي يمتلكها وذلك مستوى معنوية $\alpha = 0.05$.

-٢٧- أجري استطلاع للرأي على 300 طالب في ثلاث كليات للتجارة (في ثلاث جامعات مختلفة) وذلك لمعرفة المجال التجاري الذين يرغبون العمل فيه بعد التخرج وكانت الإجابات كالتالي :

المجال	A	B	C
الاستثمار	20	20	10
البنوك	15	25	30
الشركات	60	40	50
الاستشارات	5	15	10

يرغب المستول عن البحث في معرفة ما إذا كان المجال الذي يعمل فيه الشخص مستقل عن الجامعة التي يتخرج منها وذلك عند مستوى معنوية $\alpha = 0.05$.

-٢٨- في عينة عشوائية من 174 شخص يقودون سيارة جديدة في مدينة ما تم تصنيفهم حسب عدد السيارات التي اشتروها كل قائد سيارة خلال العشرة سنوات السابقة والأعمار المختلفة لهم وذلك في الجدول التالي :

عدد السيارات المشترقة	العمر		
	أقل من 30	30 - 40	40 فأكثر
0	49	5	12
1	23	13	12
2	11	18	7
أكثر من 2	7	12	5

قدر ما إذا كان هناك علاقة بين العمر وعدد السيارات المشتراة .

-٢٩- الجدول التالي يبين 94 شخصاً مقسمين حسب التدخين والتعليم .

التدخين	متعلم	غير متعلم
---------	-------	-----------

يدخن	29	22
لا يدخن	24	19

المطلوب اختبار العلاقة بين التدخين والتعليم عند مستوى معنوية $\alpha = 0.05$.

٣٠- تقوم إحدى الشركات بإنتاج ثلاثة مستويات من منتج معين . فإذا اختبرت عينة عشوائية مكونة من 540 وحدة من المصنع وتم تسجيل عدد الوحدات الثالثة في كل مستوى في الجدول التالي :

مستوى الإنتاج	نوع الإنتاج	
	سليم	معيب
A	185	20
B	199	24
C	98	14

اختبر فرض العدم بعدم وجود علاقة بين مستوى الإنتاج ونوع المنتج (سليم ومعيب) وذلك عند مستوى معنوية $\alpha = 0.05$.

٣١- أجرى بحث ميداني لتقدير ما إذا كان هناك علاقة بين الاتجاه السياسي وغزو الفضاء وتم الحصول على البيانات التالية :

الاتجاه السياسي	التدعيم		
	جيد جداً	جيد	لا يدعم
جمهوري	8	12	10
ديمقراطي	10	17	10
مستقل	12	6	12

اختبر فرض العدم بعدم وجود علاقة بين الاتجاه السياسي وغزو الفضاء .

٣٢- أراد مسئول في مكتبة لبيع الكتب العلمية دراسة ما إذا كان هناك علاقة بين لون الإعلانات المرسلة إلى العملاء وكمية المبيعات . أرسل 200 إعلان إلى مجموعة من العملاء وتم الحصول على البيانات التالية :

	إعلان ملون	إعلان غير ملون
الكمية المباعة	70	20
الكمية الغير مباعة	30	80

المطلوب دراسة العلاقة بين لون الإعلان وكمية المبيعات عند مستوى معنوية $\alpha = 0.05$.

٣٣- يعطى الجدول التالي التقديرات التي حصل عليها 272 طالباً في اختبارين مختلفين.

الاختبار الأول	الاختبار الثاني		
	ممتاز	جيد	مقبول
ممتاز	9	9	99
جيد	29	7	39
مقبول	59	2	19

هل يمكن الجزم بعدم وجود ارتباط بين درجات الاختبارين عند مستوى معنوية $\alpha = 0.01$.

٣٤- في عينة عشوائية من 750 شخص تم تصنيفهم حسب الدخل والوزن في الجدول التالي :

الوزن	الدخل		
	منخفض	متوسط	عالي
خفيف	100	50	50
متوسط	50	200	70
بلدين	120	60	50

أختبر فرض العدم أن العاملين (الدخل والوزن) مستقلين عند مستوى معنوية $\alpha = 0.05$.

٣٥- الجدول التالي يعطى عدد الوحدات التالفة السليمة المنتجة بواسطة آتين .

	وحدات تالفة	وحدات سليمة
الماكينة A	25	375
الماكينة B	42	558

هل عدد الوحدات التالفة مستقل عن نوع الماكينة التي تنتجها. وذلك عند مستوى معنوية

$\alpha = 0.05$.

٣٦- أعطيت لعينة من 100 شخص مسحوق للفصل لقياسي وذلك لتسجيل ما إذا كان

المسحوق فعال أو غير فعال ثم أعطى لهم مرة أخرى مسحوق جديد وتم سؤالهم مرة أخرى عن

رأيهم في المسحوق . نتائج الاستقصاء في الجدول التالي :

المسحوق القياسي	المسحوق الجديد	
	فعال	غير فعال
فعال	20	10
غير فعال	50	20

هل يمكن القول بأن هناك ارتباط بين القاعدية ونوع المسحوق ؟ وذلك عند مستوى معنوية $\alpha = 0.01$.

٣٧- يعطى الجدول التالي تصنيف لعينة عشوائية من 2764 شخص حسب الدخل بالدولار والفترة منذ آخر زيادة لاستشارة طبيب .

الدخل	منذ 6 شهور	من 7 شهور لسنة	أكثر من سنة
أقل من 3000	186	38	35
3000-4999	227	54	45
5000-6999	219	78	78
7000-9999	355	112	140
أكثر من 10,000	653	285	259

المطلوب اختبار ما إذا كان هناك استقلال بين المتغيرين (الدخل وزيارة الطبيب) وذلك عند مستوى معنوية $\alpha = 0.05$.

٣٨- قامت إحدى الشركات بعمل دورات تدريبية للعاملين بها لرفع كفاءتهم . اختبرت عينة عشوائية من 100 عامل وتم تصنيفهم تبعاً لكفاءتهم قبل وبعد عمل الدورات التدريبية والنتائج معطاة في الجدول التالي :

قبل الدورات	بعد الدورات		المجموع
	غير ماهر	ماهر	
غير ماهر	20	80	100
ماهر	2	98	100
المجموع	22	178	200

هل يمكن القول أن هناك علاقة بين التدريب والكفاءة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٣٩- قام صاحب متجر بتسجيل المعلومات التالية عن 450 شخص من المترددين على المتجر والبيانات معطاة في الجدول التالي :

	ذكور	إناث
يشترى	60	60
لا يشتري	150	180

هل يمكن القول أن الشراء يرتبط بالجنس ؟ وذلك عند مستوى معنوية $\alpha = 0.01$.

٤٠- لدراسة العلاقة بين حساسية الجلد من ضوء الشمس ولون العين حصل طبيب متخصص في الأمراض الجلدية على البيانات التالية وذلك من عينة عشوائية من 100 شخص .

لون العين	تأثير الأشعة		
	قوي	متوسط	ضعيف
أزرق	19	27	4
رمادي أو اخضر	7	8	5
بني	1	13	16

هل يمكن القول أن هناك علاقة بين لون العين وحساسية الجلد عند مستوى معنوية $\alpha = 0.05$.

٤١- أعطي دواء لثلاثة مجموعات وسجلت النتائج في الجدول التالي :

الدواء	نسبة المرضى الذين تم شفاؤهم	نسبة المرضى الذين لم يشفوا
A	95	35
B	90	10
C	85	15

قدر ما إذا كانت هذه البيانات توضح اختلاف استجابة الشفاء للأدوية الثلاثة وذلك عند مستوى معنوية $\alpha = 0.05$

٤٢- يقوم مصنعين بإنتاج نفس المنتجات A,B,C الجدول التالي يعطي الوحدات المنتجة من كل مصنع .

المصنع	المنتج			المجموع
	A	B	C	
X	42	13	33	88
Y	20	21	58	99

هل يمكن القول بأن هناك فرق معنوية بين الإنتاج للمصنعين ؟ عند مستوى معنوية $\alpha = 0.05$

٤٣- في معهد للأورام تم تجربة علاج جديد يسمى Q 134 على عينة عشوائية من 170 مريض . أيضا اختبرت عينة عشوائية من 170 مريض وتم علاجهم بالعلاج القديم المسمى Q133 والنتائج معطاة في الجدول التالي :

العلاج	شفاء	وفاة
Q134	150	20
Q133	130	40

اختبر فرض العدم $H_0: P_{j|1} = P_{j|2} = P_j; j=1,2$ وذلك عند مستوى معنوية $\alpha = 0.05$.

٤٤- اختبرت عينات عشوائية من الطلبة في فرق دراسية مختلفة وتم حساب درجاتهم وتسجيل تقديراتهم في الجدول التالي :

الفرقة	التقدير		
	ضعيف	جيد ومقبول	جيد جدا وممتاز
الفرقة الأولى	24	58	18
الفرقة الثانية	36	112	52
الفرقة الثالثة	80	230	90
الفرقة الرابعة	60	200	40

هل تدل هذه البيانات على أن توزيع الطلاب حسب تقديراتهم يعتمد على الفرقة الدراسية الموجود فيها الطالب ؟ استخدم مستوى معنوية $\alpha = 0.05$.

٤٥- يوجد في شركة سنترال A وسترال B ويعتقد العامل على السنترال A أنه يستعمل مكالمات دولية أكثر من السنترال B لاختبار هذا الفرض تم الحصول على البيانات التي في الجدول التالي :

السنترال	المكالمات المحلية	المكالمات الدولية
A	400	100
B	328	72

اختبر فرض العدم $H_0: P_{j|1} = P_{j|2} = P_j; j=1,2$ وذلك عند مستوى معنوية $\alpha = 0.05$.

٤٦- قام صاحب سيارات بعمل استطلاع للرأي عن تفضيل شراء نوع جديد من السيارات وذلك في ثلاثة مدن A, B, C. اختبرت عينات عشوائية من الثلاثة مدن وتم سؤالهم "هل تفضلون شراء السيارة الجديدة ؟" وتم الحصول على البيانات التالية :

نتائج الاستطلاع	المدن		
	A	B	C
يفضل الشراء	100	160	190
لا يفضل الشراء	50	40	60

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_j ; j = 1, 2$.

-٤٧- في استطلاع للرأي على الأشخاص بدون عمل في أربعة مناطق في بلد ما اختبرت أربع عينات من الحجم 300 واحدة من كل منطقة وتم تصنيفهم حسب المدة التي مكثوا فيها بدون عمل :

المدة بدون عمل بالأسبوع	المناطق				المجموع
	NE	SE	Nw	Sw	
أكثر من 20	80	75	60	65	280
من 20-10	115	105	110	90	420
أقل من 10	105	120	130	145	500

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_{j|4} = P_j ; j = 1, 2, 3$

-٤٨- بفرض أن جدول توافق من نوع 2×3 لمقارنة عيتين مستقلتين . بفرض أن عينة عشوائية من الرجال وعينة عشوائية أخرى من النساء أعطوا رأيهم في بحث ميداني وتم تسجيل النتائج في الجدول التالي :

	يوافق	لا يوافق	لم يقرر
سيدة	118	62	25
رجل	84	78	37

هل يمكن القول أن الرجال والسيدات يفكرون بطريقة مختلفة ؟

-٤٩- إذا تم تصنيف أنواع الفشل في طائرة إلى كهربائي - ميكانيكا - غير ذلك . فإذا تم تصميم نوعين من الطائرات I, II. المطلب اختبار فرض العدم $\alpha = 0.05$ وذلك عند مستوى معنوية $H_0 : P_{j|1} = P_{j|2} = P_j ; j = 1, 2, 3$.

وذلك بالاعتماد على البيانات التالية :

	ميكانيكي	كهربائي	غير ذلك
التصميم I	50	30	60
التصميم II	40	30	40

-٥٠- أجريت دراسة على عادة التدخين بين الذكور والإناث في منطقة ما. اختبرت عيتين واحدة من 200 رجل والأخرى من 200 سيدة والبيانات في الجدول التالي :

	يدخن	لا يدخن
رجال	110	90

سيدات	104	96
-------	-----	----

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_j, j = 1, 2$ وذلك عند مستوى معنوية

$$\alpha = 0.05$$

٥١- يعتقد شخص أن نسبة وجود صفة ما في ثلاث مجتمعات واحدة وتساوى 0.3. اختيرت عينات عشوائية من الحجم 50, 100, 50 من المجتمعات 1, 2, 3 على التوالي. يعطى الجدول التالي البيانات التي تم الحصول عليها في كل عينة. اختبر فرض العدم

$$H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_j, j = 1, 2.$$

	نعم	لا	المجموع
العينة 1	20	30	50
العينة 2	25	75	100
العينة 3	25	25	50

٥٢- بفرض أن لديك البيانات التالية في بحث ميداني :

	نعم	لا	المجموع
العينة 1	10	40	50
العينة 2	50	20	100
العينة 3	30	20	50

إذا كان يعتقد أن نسبة الذين يقولون نعم هو $P_1 = 0.8$ اختبر فرض العدم

$$H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_j, j = 1, 2$$

٥٣- في بحث ميداني تم تصنيف ثلاثة عينات من المزارعين I, II, III حسب نوع العمل في

الأرض التي يزرعها (خليط - باجر - يمتلك) . البيانات معطاة في الجدول التالي :

	يملك	باجر	خليط
العينة I	36	67	49
العينة II	31	60	49
العينة III	58	87	80

اختبر فرض العدم $H_0 : P_{j|1} = P_{j|2} = P_{j|3} = P_j, j = 1, 2, 3$ وذلك عند مستوى معنوية

$$\alpha = 0.05$$

٥٤- يتكون نظام من أربعة أجزاء تعمل مستقلة عن بعضها فإذا كان P_i يرمز إلى احتمال أن يعمل الجزء i بنجاح . أختبر فرض العدم $P_1=0.9$, $P_2=0.9$, $P_3=0.8$, $P_4=0.8$ وذلك باستخدام البيانات في الجدول التالي والتي تعطي عدد مرات النجاح لكل جزء في 50 محاولة.

الجزء	1	2	3	4
عدد مرات النجاح	40	48	45	40

٥٥- البيانات التالية تمثل 20 مفردة تم توليدها على الحاسب الآلي :

.81	.48	.1	.29	.31	.86	.91	.92	.27	.21
.31	.39	.39	.47	.84	.81	.97	.51	5.9	.70

(أ) أختبر فرض العدم $H_0 : M = 0.5$ ضد الفرض البديل $H_1 : M > 0.5$ وذلك عند

مستوى معنوية $\alpha = 0.05$.

(ب) أختبر فرض العدم $H_0 : M = 0.21$ ضد الفرض البديل $H_1 : M > 0.21$ وذلك عند

مستوى معنوية $\alpha = 0.01$.

٥٦- البيانات التالية تمثل الدخل السنوي لعشرين أسرة بالدولار

20130	25570	20410	30700	19340	2370	48160	14350	13670	5850
30700	19340	496	24840	17880	27620	21660	12110	13570	45150

أختبر فرض العدم $H_0 : M = 24800$ ضد الفرض البديل $H_1 : M < 24800$ وذلك عند

مستوى معنوية $\alpha = 0.05$.

٥٧- إذا كانت أوزان إناث القروذ بالكيلو جرام لعينة من الحجم $n=15$ هي :

8.3	9.50	9.60	8.75	8.40	9.10	9.25	9.80	10.05
8.15	10.00	9.60	9.80	9.20	9.30			

أختبر فرض العدم $M = 8.81$ ضد الفرض البديل $H_1 : M > 8.41$ وذلك باستخدام اختبار

الإشارة .

٥٨- تم سؤال 14 شخص من المدخنين عن العمر الذي بدءوا التدخين عنده وكان لدينا

النتائج الآتية :

الشخص	1	2	3	4	5	6	7
العمر	22	25	37	28	15	14	22
الشخص	8	9	10	11	12	13	14
العمر	16	18	17	23	16	20	18

هل يمكن اختيار الفرض القائل بأن وسيط المجتمع المسحوبة منه هذه العينة يساوى 20 ضد الفرض القائل أن الوسيط لا يساوى 20 وذلك عند مستوى معنوية $\alpha = 0.05$.

٥٩- إذا كانت M ترمز لوسيط توزيع متماثل من النوع المتصل استخدم اختبار الإشارة لاختبار فرض العدم $H_0: M = 75$ ضد الفرض البديل $H_1: M > 75$ وذلك عند مستوى معنوية $\alpha = 0.05$ وذلك بالاعتماد على عينة عشوائية من الحجم $n=18$. المشاهدات في الجدول التالي تعطى انحرافات مشاهدات العينة عن 75 أي $(x_i - M)$, $i=1,2,\dots,18$

1.5	-0.5	1.6	0.4	2.3	-0.8	3.2	0.9	2.9
0.3	1.8	1.5	-0.1	1.2	2.5	0.6	-0.7	1.9

٦٠- البيانات التالية تمثل أطوال عينة من 16 شجرة بالستيمترات :

122	117	127	102	137	135	124	145
130	139	127	129	143	113	116	135

أختبر ما إذا كان وسيط المجتمع هو $M = 130$ وذلك عند مستوى معنوية $\alpha = 0.05$.
(الفرض البديل $M \neq 130$).

٦١- استخدم اختبار الإشارة لتحليل فئة الأزواج في الجدول التالي وذلك لاختبار فرض العدم $H_0: M_D = 0$ ضد الفرض البديل $H_1: M_D > 0$ وذلك عند مستوى معنوية $\alpha = 0.05$.

الشخص	1	2	3	4	5	6	7	8	9	10
المعالجة A	46	41	37	32	28	43	42	51	28	27
المعالجة B	32	43	37	32	31	39	44	53	26	31

٦٢- البيانات في الجدول التالي تمثل المبيعات لمشروب بارد A في يوم ما السقي اشترها 12 شخص وأيضاً المبيعات لنفس الأشخاص من المشروب B في نفس اليوم.

المشاهدة	المشروب A	المشروب B
1	4	5
2	6	5
3	3	3
4	5	6
5	7	8
6	8	6
7	9	6
8	3	10
9	4	4
10	4	5
11	5	5

12	4	6
----	---	---

أختبر فرض العدم $H_0: M_D = 0$ ضد الفرض البديل $H_1: M_D > 0$ وذلك عند مستوى معنوية $\alpha = 0.05$.

٦٣- أخذت عينة من 10 أطفال بإحدى المدارس ودونت أوزانهم ثم أعطي كل منهم وجبة غذائية وذلك لمدة 4 شهور متتالية ثم دونت أوزانهم فكانت النتائج كالآتي :

الطفل	1	2	3	4	5	6	7	8	9	10
قبل الوجبة	130	127	129	141	137	134	139	140	138	139
بعد الوجبة	129	124	139	177	134	136	137	135	133	132

المطلوب اختبار أثر تعاطي الوجبة في زيادة وزن مجتمع هؤلاء الأطفال عند مستوى معنوية $\alpha = 0.05$.

٦٤- استخدمت طريقتين لتقدير مستوى مركب Protein Bound iodine وذلك باستخدام 10 أنش بالغة. استخدم اختبار الإشارة في تقدير ما إذا كانت النتائج في الجدول التالي والمقدرة بالطريقة B أكثر معنوية من القيم المقدرة بالطريقة A.

الشخص	1	2	3	4	5	6	7	8	9	10
A	4	6	3	5	7	8	9	3	4	5
B	5	5	3	6	8	6	10	4	5	5

٦٥- اختبر 18 زوج (توائم) من ذكور حيوان ما حيث أعطى الغذاء I لواحد من التوائم وأعطى الغذاء II للآخر في كل زوج. الأوزان خلال 8 أسابيع معطاة في الجدول التالي :

I	111	102	90	110	108	125	125	99	121
II	101	97	90	96	95	110	107	85	104
I	133	101	98	109	107	124	115	90	120
II	119	98	97	104	90	109	106	84	105

استخدم اختبار الإشارة لاختبار الفرض بعدم وجود فرق معنوي بين الغذائين ضد الفرض البديل أن الغذاء I أحسن وذلك عند مستوى معنوية $\alpha = 0.05$.

٦٦- البيانات التالية تعطي عدد ضربات القلب في الدقيقة لعشرة فئران مرة عند وضع الفأر في قفص بدون رفاق ومرة أخرى في وجود رفاق له.

الفأر	1	2	3	4	5	6	7	8	9	10
-------	---	---	---	---	---	---	---	---	---	----

وحيد	463	462	462	456	450	426	418	415	409	402
الفار مع رفاق	523	494	461	461	535	476	454	484	470	437

أخبر فرض العلم $H_0: M_D = 0$ ضد الفرض البديل $H_1: M_D \neq 0$ عند مستوى معنوية $\alpha = 0.05$.

٦٧- أجريت مسابقة لعمل كيك ذات مواصفات معينة لمجموعتين من السيدات أ و ب وكانت الدرجات التي حصلت عليها كل سيدة معطاة في الجدول التالي :

المجموعة أ	91	92	96	97	97	93	92	90
المجموعة ب	91	90	91	87	94	95	88	89

هل يمكن القول أن العنيتين تم اختيارهما من نفس المجتمع ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٦٨- يعطى الجدول التالي عدد الصمامات الكهربائية المنتجة في خطين من خطوط الإنتاج A,B وذلك خلال فترة 10 أيام .

A	170	164	140	184	174	142	191	169	161	200
B	201	179	159	195	177	170	185	179	170	212

هل يمكن القول أن العنيتين تم اختيارهما من نفس المجتمع وذلك عند مستوى معنوية $\alpha = 0.05$.

٦٩- نوعين من البلاستيك ، كل نوع منتج بطريقة مختلفة . وقد تم اختبار ultimate strength لكل مفردة في عينة عشوائية من النوع الأول وأيضاً عينة عشوائية من النوع الثاني (وحدة القياس 1,000 رطل لكل بوصة مربعة).

النوع 1	15.3	18.7	22.3	17.6	19.1	14.8
النوع 2	21.2	22.4	18.3	19.3	17.1	27.7

هل يمكن القول أن النوعين لهما نفس ultimate strength وذلك عند مستوى معنوية $\alpha = 0.05$.

٧٠- لمقارنة معدل البهض في ذكور الفئران بمعدل البهض لإناث الفئران تم الحصول على البيانات التالية :

الذكور	74	77	78	75	72	71
الإناث	60	83	73	84	82	79

هل يمكن القول أن معدل النبض واحد في الذكور والإناث وذلك عند مستوى معنوية $\alpha = 0.05$.

٧١- يعطى الجدول التالي أزمنة الحياة لكرات التجميل وذلك باستخدام اختبارين مختلفين .
المطلوب اختبار فرض العدم H_0 : المجتمعين هما نفس التوزيع ضد الفرض البديل H_1 : قيم x'_S
نتجه لأن تكون أقل من قيمة y'_S وذلك عند مستوى معنوية $\alpha = 0.05$.

الاختبار الأول	140.3	158.0	183.9	132.9	117.8	98.7	164.8	93.4
الاختبار الثاني	193.0	172.5	173.3	204.7	172.0	152.9	216.0	422.6

٧٢- مقارنة القلوية في بركتين (مقاسه بالمليجرام لكل لتر) ، أخذت عنتين من الحجم 5 من كل بركة . استخدم اختبار Wilcoxon لتقدير ما إذا كان هناك فرق معنوي في القلوية بين البركتين :

البركة I	102	116	122	112	104
البركة II	108	117	115	120	105

٧٣- إذا كانت كمية النيكوتين لنوعين من السجائر مقاس بالمليجرام كالآتي :

النوع A	22.1	24.0	26.3	24.8	25.4	26.1	23.3	22.1	24.1	22.3
النوع B	24.1	20.6	23.1	22.5	24.0	26.2	21.6	22.2	21.9	25.4

هل يمكن القول أن كمية النيكوتين واحدة في النوعين وذلك عند مستوى معنوية $\alpha = 0.05$.

٧٤- إذا كان عدد البكتريا في وحدة حجم وذلك لأربع وحدات من الحجم في المزرعة A وأيضاً في المزرعة B البيانات معطاة في الجدول التالي . اختبر فرض العدم H_0 : أن المجتمعين هما نفس التوزيع ضد الفرض البديل H_1 : قيم x'_S نتجه لأن تكون أكبر من قيم y'_S :

A	27	361	26	25
B	32	29	35	29

٧٥- استخدم اختبار Wilcoxon لتقدير ما إذا كان هناك فرق معنوي بين الفئتين التاليين من

البيانات :

A	82	86	30	21	38	29	29	19
B	124	116	54	54	110	29	39	54

٧٦- في مصنع للورق يوجد نوعين من الورق A, B. أخذت عينة عشوائية من 5 لفات مسن كل نوع وقيست قوة التمزق لمفردات كل عينة. البيانات معطاة في الجدول التالي :

النوع A	154	143	135	140	128
النوع B	149	162	160	154	175

هل يمكن القول أن قوة التمزق مختلفة في النوعين وذلك عند مستوى معنوية $\alpha = 0.05$.

٧٧- لمعرفة تأثير طريقة حديثة لزراعة الذرة استخدمت إحدى محطات التجارب الزراعية هذه الطريقة الحديثة في زراعة 8 قطع وقرون محصول الناتج مع المحصول الناتج من 9 قطع متجاورة زرعت بالطريقة العادية وكانت النتائج (عدد الشجيرات بالفدان كالآتي):

الطريقة العادية	40.1	28.7	33	36.9	34.9	33.5	37.1	40.2	
الطريقة الحديثة	55.1	49.2	23.4	46.5	52	52.2	51.3	40.9	50.1

هل هذه النتائج تعطي دليلاً على أن الطريقة الحديثة أحسن من الطريقة العادية .

٧٨- تعطي البيانات في الجدول التالي المبيعات اليومية لمحلية تجاريتين ، العينة الأولى للمحل 1

تعطي المبيعات اليومية خلال 12 يوماً والعينة الثانية تعطي المبيعات اليومية للمحل 2 خلال 10

أيام . أختبر فرض العدم H_0 : المجتمعين لهما نفس التوزيع ضد الفرض البديل H_1 : قيم x'_s تتجه لأن تكون أكبر من قيم y'_s وذلك عند مستوى معنوية $\alpha = 0.01$. المبيعات مقاسة بالدولار.

المحل 1	100	125	60	137	82	99	150	98	143	122	95	110
المحل 2	82	98	115	143	65	123	128	93	135	120		

٧٩- اختبرت مجموعتان من الأرباب الأولي من 10 أرباباً أعطيت الغذاء A والثانية مسن 12

أرباباً أعطيت غذاء B وكانت الزيادة في الوزن بعد فترة معينة هي :

A	24	34	30	30	33	22	12	23	27	29		
B	30	36	38	33	38	34	30	35	30	36	38	36

هل يمكن القول أن العيتين تم اختيارهما من نفس المجتمع ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٨٠- يعطى الجدول التالي درجات الذكاء لمجموعتين من الطلبة والمطلوب اختبار هل العيتين تم اختيارهما من نفس المجتمع وذلك عند مستوى معنوية $\alpha = 0.05$.

I	111.7	119.5	118.7	111.2	117.2	117.4
II	116.6	115.8	115.8	115.4	115.1	115.0
I	117.4	116.8	115.0	114.1	114.1	112.2
II	113.9	112.4	112.7	111.9	112.8	110.2

أختبر فرض العدم أن المجتمعين لهما نفس التوزيع ضد الفرض البديل أن قيم \bar{x}_g تتجه لأن تكون أكبر من قيم \bar{y}_g وذلك عند مستوى معنوية $\alpha = 0.05$.

٨١- البيانات التالية تعطي كمية الوقود المستهلكة لثلاث أنواع من المحركات (الوحدات المستخدمة في القياس kilometers per liter).

النوع A	31	30	29	22	32	25				
النوع B	12	13	26	24	11	26	27			
النوع C	17	20	23	9	15	18	19	14	8	5

هل يوجد فروق معنوية بين الأنواع الثلاثة عند مستوى معنوية $\alpha = 0.05$.

٨٢- يرغب مسئول في مصنع لإنتاج الورق في تقدير ما إذا كان هناك اختلاف معنوي في مقاومة الرطوبة لأربعة طرق لتخزين الورق . نتائج الاختبار لطرق التخزين الأربعة معطاة في الجدول التالي :

الطريقة			
A	B	C	D
5.3	4.3	6.0	5.6
5.3	3.7	5.0	8.0
6.5	3.8	5.6	5.4
5.4	4.6	4.9	6.5
7.6	4.1	4.5	8.5

هل يمكن القول أن هناك اختلاف في كمية الرطوبة بين طرق التخزين المختلفة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٨٣- قررت شركة تطوير ثلاثة أنواع من المحركات واختبار واحد منهم لسيارة فاخرة يسراد تسويقها . لهذا الغرض أجريت تجربة لتقدير استهلاك الوقود لكل محرك حيث اختبرت عينة

عشوائية من كل نوع وتم اختبارهم في ظروف الطريق الطبيعي . البيانات التي تم الحصول عليها من الاختبار في الجدول التالي :

1	31	30	29	22	32	25				
2	12	13	26	24	11	16	27	10		
3	3	17	20	32	9	15	18	19	14	8

المطلوب اختبار فرض العدم H_0 : توزيعات المجتمعات الثلاثة متطابقة ضد الفرض البديل H_1 :

التوزيعات ليس لها نفس الوسيط وذلك عند مستوى معنوية $\alpha = 0.01$.

٨٤- لدراسة تأثير نوع جديد من سريوم الدم على وقف سرطان الدم ، اختبر 9 فئران في مراحل متقدمة من المرض وتم تعريض 5 منهم للمعالجة و 4 لم يتم تعريضهم . يعطى الجدول التالي أزمدة الحياة ، بالسنة ، من بداية التجربة .

عولج	2.1	5.3	1.4	4.6	0.9
لم يعالج	1.9	0.5	2.8	3.1	

هل يمكن استنتاج أن سريوم الدم له تأثير فعال على إيقاف سرطان الدم ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٨٥- البيانات التالية تعطي أعمار البطاريات (بالساعة) لثلاثة أنواع من البطاريات المستخدمة في الآلات الحاسبة المحمولة .

I	II	III
12	17	21
19	25	15
19	13	16
16	31	29
20	36	32
24	15	14
30	20	18
15	35	26

هل يوجد اختلاف في متوسط أعمار البطاريات للأنواع الثلاثة ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

٨٦- في تجربة تم تقدير مستويات مركب antecubital vein cortisol لثلاث مجموعات من المرضى . البيانات معطاة في الجدول التالي :

I	262	307	211	323	454	339	304	154	287	326
II	465	501	455	355	468	360				
III	343	772	207	1048	838	687				

المطلوب اختبار فرض العدم أن توزيعات المجتمعات الثلاثة مطابقة ضد الفرض البديل أن توزيعات المجتمعات الثلاثة ليس لها نفس الوسيط وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٧- قام مسئول بتسجيل الإنتاج اليومي لثلاثة ماكينات وحصل على البيانات المعطاة في الجدول التالي :

108	95	92	110	100	الآلة الأولى
98	101	90	97	99	الآلة الثانية
103	97	113	104	98	الآلة الثالثة

أختبر فرض العدم أن الماكينات الثلاثة تعطي نفس الإنتاج في اليوم وذلك عند مستوى معنوية $\alpha = 0.05$.

-٨٨- زرعت ثلاثة أنواع من السمادة عشوائية على مجموعة من قطع الأراضي المتجاورة والمزروعة بنوع واحد من محصول النرة فكان المحصول الناتج كما يلي :

السماد الأول 83, 80, 76, 77, 85, 62

السماد الثاني 69, 50, 74, 71, 73, 68

السماد الثالث 55, 77, 65, 61, 80, 82

أختبر عند مستوى معنوية $\alpha = 0.01$ أن تأثير السماد متساوي ضد الفرض البديل أنها ليست جميعاً متساوية .

-٨٩- للمقارنة بين ثلاثة أنواع من الأدوية في إنقاص الوزن تم اختيار عينة عشوائية من 21 شخصاً وقسمت إلى ثلاثة مجموعات حيث أعطى لكل مجموعة نوع من هذا الدواء وبعد ثلاثة أشهر سجل النقص في أوزانهم بالكيلو جرام كالآتي :

الدواء

A 4.1, 5.2, 5.3, 5.6, 7.1, 7.2, 8.0

B 2.3, 2.4, 3.2, 3.2, 3.4, 3.5, 3.6

C 6.2, 6.9, 7.0, 7.8, 8.8, 9.0, 9.5

أختبر فرض العدم أن تأثير الأنواع الثلاثة متماثل ضد الفرض البديل أن هذا التأثير يختلف من نوع إلى آخر عند مستوى معنوية $\alpha = 0.05$.

-٩٠- البيانات التالية تمثل الإنتاج اليومي من الأكواب الزجاجية لثلاث ماكينات وذلك من عينة عشوائية من 12 يوماً. أختبر معنوية الفروق بين الماكينات الثلاثة وذلك عند مستوى معنوية

$\alpha = 0.05$

A 340, 345, 330, 342, 338

B 339, 333, 344

C 347, 343, 349, 355

٩١- فيما يلي مستوى مركب Protoporphyrin في الدم لـ مجموعتين من المرضى ومجموعة من الأصحاء (ملليجرام لكل 100 سم³).

مجموعة الأصحاء	21, 25, 46, 29, 37, 77, 27, 57, 71, 55, 30, 25, 45, 36, 29
المجموعة الأولى من المرضى	77, 171, 285, 81, 452, 511 173, 914, 83, 152, 770
المجموعة الثانية من المرضى	37, 28, 39, 44, 45, 28, 37, 19, 67, 11, 36, 8, 75, 147, 10

لنهل نذل هذه البيانات على اختلاف مستوى مركب Protoporphyrin في الدم لهذه المجموعات الثلاثة عند مستوى معنوية $\alpha = 0.05$.

٩٢- البيانات التالية تمثل نسبة غدة الأدرينالين لكل جرام من وزن الجسم في ثلاثة أنواع من الفئران :

المجموعة الأولى	85.4	127.79,	137.7,	139.5,	139,	155.1
المجموعة الثانية	180.3,	192.2,	190.6,	233.0,	220.5	
المجموعة الثالثة	138.6,	140.8,	144.7,	183.3,	206.6	

اختبر معنوية الفروق بين الأنواع الثلاثة عند مستوى معنوية $\alpha = 0.05$.

٩٣- في دراسة لمقارنة مستوى الكوليسترول في ثلاثة مجاميع كانت النتائج كمايلي :

المجموعة الثانية	262, 307, 200, 320, 453, 339, 303 153, 288, 355
المجموعة الثانية	464, 500, 454, 354, 467, 361
المجموعة الثالثة	342 771 206 1047 837 687

اختبر معنوية الفروق بين المجموعات الثلاثة عند مستوى معنوية $\alpha = 0.05$.

٩٤- البيانات التالية تمثل عدد الأسماك التي تم صيدها سنوياً عند موقع خاص في بحيرة كبيرة ، البيانات مقاسة بالطن . حلل هذه البيانات لتقدير ما إذا كان التغير السنوي عشوائياً أم لا ؟

السنة	1943	1944	1945	1946	1947	1948	1949	1950	1951
العدد	56	57	61	59	61	51	55	52	48
السنة	1952	1953	1954	1955	1956	1957	1958	1959	1960
العدد	51	49	47	51	41	48	39	42	41

-٩٥- البيانات التالية تعطي عدد الحمام في منطقة معينة (محمية) خلال السنوات من 1954 حتى 1965 .

السنة	1954	1955	1956	1957	1958	1959
العدد	13	13	12	10	8	11
السنة	1960	1961	1962	1963	1964	1965
العدد	9	7	6	5	7	8

هل تعتقد أن حجم المجتمع يتأثر عشوائياً من سنة إلى أخرى ؟ وذلك عند مستوى معنوية $\alpha = 0.05$.

-٩٦- البيانات التالية تمثل الأزمنة بين الحوادث في مدينة كبيرة
8.66 11.28, 10.43, 10.89, 11.49, 15.92, 12.5, 13.86, 13.32
أختبر فرض العدم أن العينة عشوائية ضد الفرض البديل أن العينة غير عشوائية عند مستوى معنوية $\alpha = 0.05$.

-٩٧- بفرض أن لدينا عينة من 25 شخص وكنا نرمز للشخص الذكر في العينة بالرمز M وبالرمز F للإناث وكانت لدينا النتائج الآتية
MF MMMM F MMM FFFF M FF MMM FFFFF
المطلوب عند مستوى معنوية $\alpha = 0.05$ تحديد هل العينة عشوائية أم لا ؟

-٩٨- يعطى الجدول التالي درجات حرارة موجبة وسالبة
9, 10, 11.5, 3, -1, 3, 4, -7, -6, 3,
-9, 5, 9, 6, 5, 2, 7, -2, -5, -4, -6
والمطلوب بمستوى معنوية $\alpha = 0.05$ معرفة هل هذه البيانات عشوائية بالنسبة للإشارة أم لا ؟
-٩٩- أجرى اختبار مجموعة من الموظفين قبل تدريبهم وبعد 6 أشهر من التدريب وقد تم ترتيبهم تبعاً لجودة إنتاجهم . البيانات التالية تعطي الرتب التي حصلوا عليها .

الموظف	A	B	C	D	E	F	G	H	I	J	K	L
الرتب قبل التدريب	2	3	1	6	7	4	8	5	12	11	9	10
الرتب بعد التدريب	1	2	3	4	5	6	7	8	9	10	11	12

أوجد معامل سبيرمان .

-١٠٠- في وكالة لبيع السيارات أجريت دراسة على 15 موظف في قسم المبيعات وذلك لدراسة العلاقة بين درجة الاختبار التي حصل عليها الموظف عند تعيينه وعدد السيارات المباعة خلال السنة الأولى من التعيين

الشخص	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
الدرجة	72	88.5	70	87	71	85	89	93	98	96	86	82	88	83	80
عدد السيارات	314	422	322	440	287	415	463	497	510	512	432	390	453	374	385

أختبر فرض العدم H_0 : المتغيرين مستقلين ضد الفرض البديل H_1 : توجد علاقة بين المتغيرين في نفس الاتجاه وذلك عند مستوى معنوية $\alpha = 0.05$.

-١٠١- أجريت مقابلة شخصية من قبل شخصين A, B وذلك لمائة أشخاص متقدمين لوظيفة ما وتم تسجيل الدرجات التالية :

الشخص	1	2	3	4	5	6	7	8
A	12	17	14	7	10	12	6	3
B	16	15	18	5	3	10	7	4

حول الدرجات إلى رتب وأحسب معامل سيرمان .

-١٠٢- البيانات التالية تعطي الرتب التي حصل عليها 10 سكرتيرات يكتبن على الآلة الكاتبة أولاً في الظروف العادية ثم في ظروف اختبار . أوجد معامل سيرمان

السكرتيرة	A	B	C	D	E	F	G	H	I	J
في ظروف الاختبار	1	2	3	4	5	6	7	8	9	10
في الظروف العادية	3	4	6	1	2	8	9	10	7	5

-١٠٣- قامت شركة بإجراء اختبار لـ 7 أشخاص من المتقدمين لوظيفة ما . الاختبار الأول مقابلة شخصية والآخر تحريري وتم الحصول على البيانات التالية :

الشخص	1	2	3	4	5	6	7
الرتبة للمقابلة الشخصية	4	1	7	6	2	3	5
الرتبة للامتحان التحريري	5	2	7	4	1	3	6

أوجد معامل سيرمان .

-١٠٤- طلب أستاذ في الجامعة من 10 طلاب تقديم بحثين في الامتحانات النهائي . الدرجات التي حصل عليها كل طالب هي :

الطالب	1	2	3	4	5	6	7	8	9	10
درجة البحث الأول (الدرجة من 100)	58	56	54	65	58	60	59	51	53	56
درجة البحث الثاني (الدرجة من 10)	6	7	5	8	10	5	7	6	2	4

أوجد معامل سيرمان .

-١٠٥- الجدول التالي يسجل كمية المطر (اليومية) وعدد ساعات ظهور السحب (بالساعات) في بلد ما خلال 11 فترة زمنية والمطلوب إيجاد معامل سيرمان للرتب بين سقوط الأمطار وظهور السحب.

الفترة	1	2	3	4	5	6	7	8	9	10	11
سقوط المطر	15.1	15.8	14.9	16.6	12.6	17.4	16.1	13.7	15.5	17.5	13.2
ظهور السحب	1692	1634	1835	1741	1876	1561	1921	1942	1822	1542	1874

-١٠٦- في روضة للأطفال اختبرت عينة عشوائية من 20 طفل وتم تسجيل درجة كل طفل في الامتحان وكذلك عمره والبيانات معطاة في الجدول التالي :

العمر	6.50	6.75	7.0	7.50	7.50	7.50	7.50	7.75	8.0	8.00
الدرجة	16	28	46	14	41	10	56	43	15	21
العمر	8.25	8.25	8.75	9.00	9.75	9.50	9.50	9.75	9.00	9.00
الدرجة	28	57	36	71	47	66	71	61	60	72

استخدم سيرمان لاختبار العلاقة بين العمر والدرجة عند مستوى معنوية $\alpha = 0.5$.

REFERENCES

المراجع

أولاً : المراجع العربية

- ١- أحمد عبادة سرحان، (١٩٦٨)، مقدمة في طرق التحليل الإحصائي، معهد الدراسات والبحوث الإحصائية-القاهرة .
- ٢- أحمد عبادة سرحان، (١٩٨٣)، تصميم التحارب وتحليلها، دار الكتب الجامعية-القاهرة- جمهورية مصر العربية .
- ٣- أحمد فؤاد غالب وآخرون، (١٩٩٣)، الرياضة للدراسي العلوم الحيوية-ترجمة لكتاب جاديش س. أريا و روبين و. لاردنر-الدار الدولية للنشر والتوزيع-القاهرة-جمهورية مصر العربية .
- ٤- أنيس إسماعيل كنحو، (١٩٩٣)، الإحصاء والاحتمال-جامعة الملك سعود-عمادة شؤون المكتبات .
- ٥- بدرية شوقي عبد الوهاب و محمد كامل الشربينى، (١٩٨٤)، المبادئ الأولية في الإحصاء -ترجمة لكتاب بول ج. هويل-الطبعة الرابعة-دار جون وايلي وأبنائه .
- ٦- جلال الصياد، (١٩٨٨)، نظرية الاحتمالات-الطبعة الثانية-دار الشروق-جدة-المملكة العربية السعودية .
- ٧- جلال مصطفى الصياد و محمد الدسوقي حبيب، (١٩٩٠)، مقدمة في الطرق الإحصائية -الطبعة الثانية-قائمة-جدة-المملكة العربية السعودية .
- ٨- ذكرى الشربينى، (١٩٩٥)، الإحصاء و تصميم التحارب في بحوث النفسية والتربوية و الاجتماعية-مكتبة الأنجلو المصرية-القاهرة .
- ٩- ربيع ذكى عامر، (١٩٨٩)، تحليل الانحدار-أساليبه و تطبيقاته العملية باستخدام البرنامج الجاهز SPSS/PC+ -معهد الدراسات والبحوث الإحصائية-جامعة القاهرة .
- ١٠- سعدية حافظ متصر، (١٩٨٢)، ملخصات شوم-نظريات و مسائل في الإحصاء و الاقتصاد القياسي-ترجمة لكتاب دومينيك سالفاتور-دار ماكجروهيل-نيويورك .
- ١١- سمير كامل عاشور و سامية سالم أبو الفتوح، (١٩٩٠)، مقدمة في الإحصاء التحليلي-معهد الدراسات والبحوث الإحصائية-جامعة القاهرة .
- ١٢- سمير كامل عاشور و سامية سالم أبو الفتوح، (١٩٩٠)، مقدمة في الإحصاء الوصفي-معهد الدراسات والبحوث الإحصائية-جامعة القاهرة .

- ١٣-سمير كامل عاشور و سامية سالم أبو الفتوح، (١٩٩٥)، الاختبارات اللامعلمية-معهد الدراسات والبحوث الإحصائية-جامعة القاهرة .
- ١٤-عدنان بن ماحد عبد الرحمن برى و محمود محمد إبراهيم هندي و أنور أحمد محمد عبد الله، (١٩٩١)، مبادئ الإحصاء و الاحتمالات-عمادة شؤون المكتبات-جامعة الملك سعود-المملكة العربية السعودية .
- ١٥-عفاف الدش، (١٩٩٤)، الإحصاء التطبيقي للتجارين-الطبعة الثانية-جامعة حلوان-القاهرة .
- محمد صبحي أبو صالح و عدنان محمد عوض، (١٩٨٣)، مقدمة في الإحصاء-الطبعة الرابعة-دار جون وابلبي و أبنائه-نيويورك .
- ١٧-محمد محمد الطاهر الإمام، (١٩٩٤)، تصميم و تحليل التجارب-دار المريخ-الرياض-المملكة العربية السعودية .

ثانياً : المراجع الأجنبية

- 1-Bain, L. J. (1992) Introduction to Probability and Mathematical Statistics, Second Edition, Duxbury Press - An Imprint of Wadsworth Publishing Company Belmont, California.
- 2-Cangelosi, V. E.; Taylor, P. H. and Rice, P. F. (1979) Basic statistics - A Real World Approach, Second Edition, West Publishing Company, New York.
- 3-Cochran, W. G. (1963) Sampling Techniques, Second Edition, New York : John Wiley & Sons, Inc.
- 3-Daniel, W. W. (1978) Applied Nonparametric Statistic, Houghton Mifflin Company, London.
- 4-Devore, J. L. (1995) Probability and Statistics for Engineering and the Sciences, Fourth Edition, Duxbury Press-An International Thomson Publishing Company, London.
- 5-Draper, N. R. and Smith, H. (1981) Applied Regression Analysis, Second Edition, John Wiley & Sons Inc., U.S.A.
- 6-Frank, H. and Althoen, S. C. (1997) Statistics- Concepts and Applications, Low Price Edition, Cambridge University Press.
- 7-Freund, J. E. and Williams, F. J. (1972) Elementary Business Statistics, Englewood Cliffs, N. J. Prentice-Hall, Inc.
- 8-Hamburg, M. (1979) Basic Statistics : A Modern Approach, Second Edition, Harcourt Brace Jovanovich, Inc., New York.
- 9-Mendenhall, W. (1975) Introduction to Probability and Statistics, Company, Inc. Belmont, California Fourth Edition, Duxbury Press, A Division of Wadsworth Publishing Company, Inc. Belmont, California.
- 10-Neter, J.; Wasserman, W. and Whitmore, G. A. (1993) Applied Statistics, Fourth Edition, ALLYN AND BACON, London.

- 11-Owen, F. and Jones, R. (1994) Statistics, Fourth Edition, Pitman Publishing, London.
- 12-Scheffler, W. (1979) Statistics for the Biological Sciences, Second Edition, Addison-Wesley Publishing Company. Inc. Philippines.
- 13-Walpole, R. E. (1982) Introduction to Statistics, Macmillan Publishing Co. Inc. New York.
- 14-Weisberg, S. (1980), Applied Linear Regression, John Wiley & Sons Inc., New York, U.S.A.
- 15-Winer, B. J. Brown, D. R. and Michels, K. M.(1991) Statistical Experimental Design, Third Edition, McGraw-Hill, Inc., New York.
- 16-Yamane, T. (1967) Elementary Sampling Theory, Prentice-Hall, Inc., Englewood Cliffs, N. J.
- 17-Yates, F. (1934) Contingency Tables Involving Small Numbers and the χ^2 Test, J. Roy. Statist. Soc., 1,217-235.

الملاحق

ملحق (١) جدول حساب $\sum_{x=0}^{\infty} b(x; n, p)$ لتغير عشوائي يتبع توزيع ذي الحدين .

ملحق (٢) جدول حساب $\sum_{x=0}^{\infty} p(x; \mu)$ لتغير عشوائي يتبع توزيع بواسون .

ملحق (٣) جدول المساحات تحت المنحنى الطبيعي القياسي $P(0 < Z < z)$.

ملحق (٤) جدول القيم الحرجة t_{α} لتوزيع t .

ملحق (٥) جدول القيم الحرجة χ_{α}^2 لتوزيع χ^2 .

ملحق (٦) جدول القيم الحرجة $f_{\alpha}(v_1, v_2)$ لتوزيع F عند $(\alpha = 0.05)$.

ملحق (٧) جدول القيم الحرجة $f_{\alpha}(v_1, v_2)$ لتوزيع F عند $(\alpha = 0.01)$.

ملحق (٨) جدول القيم الحرجة لاختبار Cochran χ^2 لتجانس التباين .

ملحق (٩) جدول القيم الحرجة $q_{\alpha}(p, v)$ لـ q لـ p .

ملحق (١٠) جدول القيم الحرجة c_{α} لاختبار الاعتدال .

ملحق (١١) جدول القيم الحرجة $d(n, \alpha'), d(n, \alpha'')$ لاختبار إشارة الرب .

ملحق (١٢) جدول القيم الحرجة لاختبار Mann-Whitney-Wilcoxon

ملحق (١٣) جدول القيم الحرجة لاختبار Kruskal - Wallis .

ملحق (١٤) جدول القيم الحرجة F_1 السفلي لاختبار الدورات .

ملحق (١٥) جدول القيم الحرجة F_2 العليا لاختبار الدورات .

ملحق (١٦) جدول القيم الحرجة F_{α} لاختبار سيرمان .

جدول حساب $\sum_{x=0}^r b(x; n, p)$ لمتغير عشوائي يتبع توزيع ذي الحدين

a. n=5

		p														
r		0.01	0.05	0.1	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
	0	.951	.774	.590	.328	.237	.168	.078	.031	.010	.002	.001	.000	.000	.000	.000
	1	.999	.977	.919	.737	.633	.528	.337	.188	.087	.031	.016	.007	.000	.000	.000
	2	1.00	.999	.991	.942	.896	.837	.683	.500	.317	.163	.084	.038	.009	.001	.000
	3	1.00	1.00	1.00	.993	.984	.969	.913	.812	.663	.472	.367	.263	.081	.023	.001
	4	1.00	1.00	1.00	1.00	.999	.998	.990	.969	.922	.832	.763	.672	.410	.226	.049

b. n=10

		p														
r		0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
	0	.904	.599	.349	.107	.056	.028	.006	.001	.000	.000	.000	.000	.000	.000	.000
	1	.996	.914	.736	.376	.244	.149	.046	.011	.002	.000	.000	.000	.000	.000	.000
	2	1.00	.988	.930	.678	.526	.383	.167	.055	.012	.002	.000	.000	.000	.000	.000
	3	1.00	.999	.987	.879	.776	.650	.382	.172	.055	.011	.004	.001	.000	.000	.000
	4	1.00	1.00	.998	.967	.922	.850	.633	.377	.166	.047	.020	.006	.000	.000	.000
	5	1.00	1.00	1.00	.994	.980	.953	.834	.623	.367	.150	.078	.033	.002	.000	.000
	6	1.00	1.00	1.00	.999	.996	.989	.945	.828	.618	.350	.224	.121	.013	.001	.000
	7	1.00	1.00	1.00	1.00	1.00	.998	.988	.945	.833	.617	.474	.322	.070	.012	.000
	8	1.00	1.00	1.00	1.00	1.00	1.00	.998	.989	.954	.851	.756	.624	.264	.086	.004
	9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.994	.972	.944	.893	.651	.401	.096

تابع ملحق (١)

جدول حساب $\sum_{x=0}^r b(x; n, p)$ لمطير عشوائي يتبع توزيع ذي الحدين

c. n=15

		p														
r		0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
	0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
	2	1.00	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
	3	1.00	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
	4	1.00	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
	5	1.00	1.00	.998	.939	.852	.722	.403	.151	.034	.004	.001	.000	.000	.000	.000
	6	1.00	1.00	1.00	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
	7	1.00	1.00	1.00	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
	8	1.00	1.00	1.00	.999	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
	9	1.00	1.00	1.00	1.00	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000
	10	1.00	1.00	1.00	1.00	1.00	.999	.991	.941	.783	.485	.314	.164	.013	.001	.000
	11	1.00	1.00	1.00	1.00	1.00	1.00	.998	.982	.909	.703	.539	.352	.056	.005	.000
	12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.996	.973	.873	.764	.602	.184	.036	.000
	13	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.995	.965	.920	.833	.451	.171	.010
	14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.995	.987	.965	.794	.537	.140

المصدر : عن [Devore(1995)]

تابع ملحق (١)

جدول حساب $\sum_{x=0}^r b(x; n, p)$ لتغير عشوائي يتبع توزيع ذي الحدين

d. n=20

		p														
r		0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
	0	.818	.358	.122	.012	.003	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.983	.736	.392	.069	.024	.008	.001	.000	.000	.000	.000	.000	.000	.000	.000
	2	.999	.925	.677	.206	.091	.035	.004	.000	.000	.000	.000	.000	.000	.000	.000
	3	1.00	.984	.867	.411	.225	.107	.016	.001	.000	.000	.000	.000	.000	.000	.000
	4	1.00	.997	.957	.630	.415	.238	.051	.006	.000	.000	.000	.000	.000	.000	.000
	5	1.00	1.00	.989	.804	.617	.416	.126	.021	.002	.000	.000	.000	.000	.000	.000
	6	1.00	1.00	.998	.913	.786	.608	.250	.058	.006	.000	.000	.000	.000	.000	.000
	7	1.00	1.00	1.00	.968	.898	.772	.416	.132	.021	.001	.000	.000	.000	.000	.000
	8	1.00	1.00	1.00	.990	.959	.887	.596	.252	.057	.005	.001	.000	.000	.000	.000
	9	1.00	1.00	1.00	.997	.986	.952	.755	.412	.128	.017	.004	.001	.000	.000	.000
	10	1.00	1.00	1.00	.999	.996	.983	.872	.588	.245	.048	.014	.003	.000	.000	.000
	11	1.00	1.00	1.00	1.00	.999	.995	.943	.748	.404	.113	.041	.010	.000	.000	.000
	12	1.00	1.00	1.00	1.00	1.00	.999	.979	.868	.584	.228	.102	.032	.000	.000	.000
	13	1.00	1.00	1.00	1.00	1.00	1.00	.994	.942	.750	.392	.214	.087	.002	.000	.000
	14	1.00	1.00	1.00	1.00	1.00	1.00	.998	.979	.874	.584	.383	.196	.011	.000	.000
	15	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.994	.949	.762	.585	.370	.043	.003	.000
	16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.984	.893	.775	.589	.133	.016	.000
	17	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.996	.965	.909	.794	.323	.075	.001
	18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.992	.976	.931	.608	.264	.017
	19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.997	.988	.878	.642	.182

تابع ملحق (١) جدول حساب $\sum_{x=0}^r b(x; n, p)$ لتغير عشوائي يتبع توزيع ذي الحدين

d. n=25

		p														
r		0.01	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95	0.99
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	0	.778	.277	.072	.004	.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.974	.642	.271	.027	.007	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000
	2	.998	.873	.537	.098	.032	.009	.000	.000	.000	.000	.000	.000	.000	.000	.000
	3	1.00	.966	.764	.234	.096	.033	.002	.000	.00	.000	.000	.000	.000	.000	.000
	4	1.00	.993	.902	.421	.214	.090	.009	.000	.000	.000	.000	.000	.000	.000	.000
	5	1.00	.999	.967	.617	.378	.193	.029	.002	.000	.000	.000	.000	.000	.000	.000
	6	1.00	1.00	.991	.780	.561	.341	.074	.007	.000	.000	.000	.000	.000	.000	.000
	7	1.00	1.00	.998	.891	.727	.512	.154	.022	.001	.000	.000	.000	.000	.000	.000
	8	1.00	1.00	1.00	.953	.851	.677	.274	.054	.004	.000	.000	.000	.000	.000	.000
	9	1.00	1.00	1.00	.983	.929	.811	.425	.115	.013	.000	.000	.000	.000	.000	.000
	10	1.00	1.00	1.00	.994	.970	.902	.586	.212	.034	.002	.000	.000	.000	.000	.000
	11	1.00	1.00	1.00	.998	.980	.956	.732	.345	.078	.006	.001	.000	.000	.000	.000
	12	1.00	1.00	1.00	1.00	.997	.983	.846	.500	.154	.017	.003	.000	.000	.000	.000
	13	1.00	1.00	1.00	1.00	.999	.994	.922	.655	.268	.044	.020	.002	.000	.000	.000
	14	1.00	1.00	1.00	1.00	1.00	.998	.966	.788	.414	.098	.030	.006	.000	.000	.000
	15	1.00	1.00	1.00	1.00	1.00	1.00	.987	.885	.575	.189	.071	.017	.000	.000	.000
	16	1.00	1.00	1.00	1.00	1.00	1.00	.996	.946	.726	.323	.149	.047	.000	.000	.000
	17	1.00	1.00	1.00	1.00	1.00	1.00	.999	.978	.846	.488	.273	.109	.002	.000	.000
	18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.993	.926	.659	.439	.220	.009	.000	.000
	19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998	.971	.807	.622	.383	.033	.001	.000
	20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.991	.910	.786	.579	.098	.007	.000
	21	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998	.967	.904	.766	.236	.034	.000
	22	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.991	.968	.902	.463	.127	.002
	23	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998	.993	.973	.729	.358	.026
	24	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.996	.928	.723	.222

جدول حساب $\sum_{x=0}^r p(x; \mu)$ لمغير عشوائي يتبع توزيع بواسون

		μ									
r		.1	.2	.3	4	5	.6	.7	.8	.9	1.0
	0	.905	.819	.741	.670	.607	.549	.497	.449	.407	.368
	1	.995	.982	.963	.938	.910	.878	.844	.809	.772	.736
	2	1.00	.999	.996	.992	.986	.977	.966	.953	.937	.920
	3		1.00	1.00	.999	.998	.997	.994	.991	.987	.981
	4				1.00	1.00	1.00	.999	.999	.998	.996
	5							1.00	1.00	1.00	.999
	6										1.00

المصدر عن : [Devore (1995)]

[illegible]

ملحق (٣)

جدول المساحات تحت المنحنى الطبيعي القياسي

$$P(0 < Z < z)$$

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

المصدر : عن [Daniel (1978)]

ملحق (٤)

جدول القيم الحرجة t_{α} لتوزيع t

v	α					
	.10	.05	.025	.01	.005	.0005
1	3.078	6.314	12.706	31.821	63.657	318.31
2	1.886	2.920	4.303	6.965	9.925	22.326
3	1.638	2.353	3.182	4.541	5.841	10.213
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.878	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
60	1.296	1.671	2.000	2.390	2.660	3.232
120	1.289	1.658	1.980	2.358	2.617	3.160
∞	1.282	1.645	1.960	2.326	2.576	3.090

المصدر : عن [Devore (1995)]

جدول القيم الحرجة χ^2_{α} لتوزيع χ^2

V	α									
	.995	.99	.975	.95	.90	.10	.05	.025	.01	.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.523	13.120	14.611	16.473	34.381	37.652	40.646	44.313	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.708	19.768	39.087	42.557	45.772	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

المصدر : عن [Devore(1995)]

ملحق (٦)

جدول القيم الحرجة $f_{\alpha}(v_1, v_2)$ لتوزيع F عند $(\alpha = 0.05)$

v_2		v_1																		
		1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	
1	2	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	3	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	4	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.636
5	6	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	7	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	8	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	9	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	10	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	11	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	12	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	13	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	14	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	15	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	16	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	17	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.02
17	18	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	19	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	20	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	21	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	22	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	23	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	24	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	25	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	26	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	27	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.33	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	28	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	29	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	30	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	40	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	60	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	120	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	∞	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞		3.84	3.84	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

المصدر : عن [Devore (1995)]

(V) ملحق

جدول القيم الحرجة $f_{\alpha}(v_1, v_2)$ لتوزيع F عند $\alpha = 0.01$

v_2	v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1		4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2		98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4		21.20	18.00	16.69	15.96	15.52	15.21	14.90	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5		16.16	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6		13.57	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7		12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.71	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8		11.24	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9		10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10		10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11		9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12		9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.369
13		9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14		8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16		8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17		8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18		8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19		8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20		8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21		8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22		7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23		7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24		7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25		7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26		7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27		7.66	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28		7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29		7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30		7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40		7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60		7.08	4.96	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120		6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.55	1.38
∞		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

جدول القيم الحرجة لاختبار Cochran لتجانس التباين

df for S_j^2	$1-\alpha$	k=number of variances										
		2	3	4	5	6	7	8	9	10	15	20
1	.95	.9985	.9669	.9045	.8412	.7888	.7271	.6798	.6385	.6020	.4709	.3894
	.99	.9999	.9933	.9676	.9279	.8828	.8376	.7945	.7544	.7175	.5747	.4799
2	.95	.9780	.8709	.7679	.6838	.6161	.5612	.5157	.4775	.4450	.3346	.2705
	.99	.9950	.9423	.8643	.7885	.7218	.6644	.6152	.5727	.5358	.4069	.3297
3	.95	.9392	.7977	.6841	.5981	.5321	.4800	.4377	.4027	.3733	.2758	.2205
	.99	.9794	.8831	.7814	.6957	.6258	.5685	.5209	.4810	.4469	.3317	.2654
4	.95	.9057	.7457	.6287	.5441	.4803	.4307	.3910	.3584	.3311	.2419	.1921
	.99	.9586	.8335	.7212	.6329	.5635	.5080	.4627	.4251	.3934	.2882	.2288
5	.95	.8772	.7071	.5895	.5065	.4447	.3974	.3595	.3286	.3029	.2195	.1735
	.99	.9373	.7933	.6761	.5875	.5195	.4659	.4226	.3870	.3572	.2593	.2048
6	.95	.8534	.6771	.5698	.4783	.4184	.3726	.3362	.3067	.2823	.2034	.1602
	.99	.9172	.7606	.6410	.5531	.4866	.4347	.3932	.3592	.3308	.2386	.1877
7	.95	.8332	.6530	.5365	.4564	.3980	.3535	.3185	.2901	.2666	.1911	.1501
	.99	.8988	.7335	.6129	.5259	.4608	.4105	.3704	.3378	.3106	.2228	.1748
8	.95	.8159	.6333	.5175	.4387	.3817	.3384	.3043	.2768	.2541	.1815	.1422
	.99	.8823	.7107	.5897	.5037	.4401	.3911	.3522	.3207	.2945	.2104	.1646
9	.95	.8010	.6167	.5017	.4241	.3682	.3259	.2926	.2659	.2439	.1736	.1357
	.99	.8674	.6912	.5702	.4854	.4229	.3751	.3373	.3067	.2813	.2002	.1567
16	.95	.7341	.5466	.4366	.3645	.3135	.2756	.2462	.2226	.2032	.1429	.1108
	.99	.7949	.6059	.4884	.4094	.3529	.3105	.2779	.2514	.2297	.1612	.1248
36	.95	.6602	.4748	.3720	.3066	.2612	.2278	.2022	.1820	.1655	.1144	.0879
	.99	.7067	.5153	.4057	.3351	.2858	.2494	.2214	.1992	.1811	.1251	.0960
144	.95	.5813	.4031	.3093	.2513	.2119	.1833	.1616	.1446	.1308	.0889	.0675
	.99	.6062	.4230	.3251	.2644	.2229	.1929	.1700	.1521	.1376	.0934	.0709

المصدر : [Winer et al (1991)]

لبنون

جدول القيم الحرجة $q_{\alpha}(p, v)$ للتباين

V	Q_L	2	3	4	5	6	7	8	9	10	11
5	.05	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17
	.01	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48
6	.05	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65
	.01	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30
7	.05	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30
	.01	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55
8	.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05
	.01	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03
9	.05	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87
	.01	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65
10	.05	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72
	.01	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36
11	.05	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61
	.01	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13
12	.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51
	.01	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94
13	.05	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43
	.01	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79
14	.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36
	.01	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66
15	.05	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31
	.01	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55
16	.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26
	.01	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46
17	.05	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21
	.01	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38
18	.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17
	.01	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31
19	.05	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14
	.01	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14	6.25
20	.05	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11
	.01	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19
24	.05	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01
	.01	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02
30	.05	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.91
	.01	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85
40	.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82
	.01	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69
60	.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73
	.01	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53
120	.05	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64
	.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37
	.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55
	.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23

لعمري

جدول القيم المخرجة $q_{\alpha}(p, v)$ للماتكن

III										Q	V
12	13	14	15	16	17	18	19	20			
7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21	.05	5	
10.70	10.89	11.00	11.14	11.40	11.55	11.68	11.81	11.93	.01		
6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59	.05	6	
9.48	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54	.01		
6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.10	7.17	.05	7	
8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65	.01		
6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87	.05	8	
8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03	.01		
5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64	.05	9	
7.78	7.91	8.03	8.13	8.23	8.33	8.41	8.49	8.57	.01		
5.83	5.93	6.03	6.11	6.19	6.27	6.34	6.40	6.47	.05	10	
7.49	7.60	7.71	7.81	7.91	7.99	8.08	8.15	8.23	.01		
5.71	5.81	5.90	5.98	6.06	6.13	6.20	6.27	6.33	.05	11	
7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95	.01		
5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21	.05	12	
7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73	.01		
5.53	5.63	5.71	5.79	5.86	5.93	5.99	6.05	6.11	.05	13	
6.90	7.01	7.10	7.19	7.27	7.35	7.42	7.48	7.55	.01		
5.46	5.55	5.64	5.71	5.79	5.85	5.91	5.97	6.03	.05	14	
6.77	6.87	6.96	7.06	7.13	7.20	7.27	7.33	7.39	.01		
5.40	5.49	5.57	5.65	5.72	5.78	5.85	5.90	5.96	.05	15	
6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26	.01		
5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90	.05	16	
6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15	.01		
5.31	5.39	5.47	5.54	5.61	5.67	5.73	5.79	5.84	.05	17	
6.48	6.57	6.66	6.73	6.81	6.87	6.94	7.00	7.05	.01		
5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79	.05	18	
6.41	6.50	6.58	6.65	6.73	6.79	6.85	6.91	6.97	.01		
5.23	5.31	5.39	5.46	5.53	5.59	5.65	5.70	5.75	.05	19	
6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89	.01		
5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71	.05	20	
6.28	6.37	6.45	6.52	6.59	6.66	6.71	6.77	6.82	.01		
5.10	5.18	5.25	5.32	5.38	5.44	5.49	5.55	5.59	.05	24	
6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61	.01		
5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.47	.05	30	
5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41	.01		
4.90	4.98	5.04	5.11	5.16	5.22	5.27	5.31	5.36	.05	40	
5.76	5.83	5.90	5.96	6.02	6.07	6.12	6.16	6.21	.01		
4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24	.05	60	
5.60	5.67	5.73	5.78	5.84	5.89	5.93	5.97	6.01	.01		
4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13	.05	120	
5.44	5.50	5.56	5.61	5.66	5.71	5.75	5.79	5.83	.01		
4.62	4.68	4.74	4.80	4.85	4.89	4.93	4.97	5.01	.05	∞	
5.39	5.45	5.50	5.55	5.60	5.64	5.68	5.72	5.76	.01		

ملحق (١٠)

جدول القيم الحرجة c_{α} لاختبار الإعتدال

		α		
		.1	.05	.01
n	5	.9033	.8804	.8320
	10	.9347	.9180	.8804
	15	.9506	.9383	.9110
	20	.9600	.9503	.9290
	25	.9662	.9582	.9408
	30	.9707	.9639	.9490
	40	.9767	.9715	.9597
	50	.9807	.9764	.9664
	60	.9835	.9799	.9710
	75	.9865	.9835	.9757

المصدر عن : [Daniel (1978)]

جدول القيم الحرجة $d(n, \alpha'')$, $d(n, \alpha')$ لاختبار إشارة الرتب

n	d	Confidence coefficient	α''	α'
3	1	.750	.250	.125
4	1	.875	.125	.063
5	1	.938	.062	.031
	2	.875	.125	.063
6	1	.969	.031	.016
	2	.937	.063	.031
	3	.906	.094	.047
	4	.844	.156	.078
7	1	.984	.016	.008
	2	.969	.031	.016
	4	.922	.078	.039
	5	.891	.109	.055
8	1	.992	.008	.004
	2	.984	.016	.008
	4	.961	.039	.020
	5	.945	.055	.027
	6	.922	.078	.039
	7	.891	.109	.055
9	2	.992	.008	.004
	3	.988	.012	.006
	6	.961	.039	.020
	7	.945	.055	.027
	9	.902	.098	.049
	10	.871	.129	.065
10	4	.990	.010	.005
	5	.986	.014	.007
	9	.951	.049	.024
	10	.936	.064	.032
	11	.916	.084	.042
	12	.895	.105	.053
11	6	.990	.010	.005
	7	.986	.014	.007
	11	.958	.042	.021
	12	.946	.054	.027
	14	.917	.083	.042
	15	.898	.102	.051
12	8	.991	.009	.005
	9	.988	.012	.006
	14	.958	.042	.021
	15	.948	.052	.026
	18	.908	.092	.046
	19	.890	.110	.055
13	10	.992	.008	.004
	11	.990	.010	.005
	18	.952	.048	.024
	19	.943	.057	.029
	22	.906	.094	.047
	23	.890	.110	.055

تابع : ملحق (١١)

جدول القيم الحرجة $d(n, \alpha''), d(n, \alpha')$ لاختبار إشارة الرتب

n	d	Confidence coefficient	α''	α'
14	13	.991	.009	.004
	14	.989	.011	.005
	22	.951	.049	.025
	23	.942	.058	.029
	26	.909	.091	.045
	27	.896	.104	.052
15	16	.992	.008	.004
	17	.990	.010	.005
	26	.952	.048	.024
	27	.945	.055	.028
	31	.905	.095	.047
	32	.893	.107	.054
16	20	.991	.009	.005
	21	.989	.011	.006
	30	.956	.044	.022
	31	.949	.051	.025
	36	.907	.093	.047
	37	.895	.105	.052
17	24	.991	.009	.005
	25	.989	.011	.006
	35	.955	.045	.022
	36	.949	.051	.025
	42	.902	.098	.049
	43	.891	.109	.054
18	28	.991	.009	.005
	29	.990	.010	.005
	41	.952	.048	.024
	42	.946	.054	.027
	48	.901	.099	.049
	49	.892	.108	.054
19	33	.991	.009	.005
	34	.989	.011	.005
	47	.951	.049	.025
	48	.945	.055	.027
	54	.904	.096	.048
	55	.896	.104	.052
20	38	.991	.009	.005
	39	.989	.011	.005
	53	.952	.048	.024
	54	.947	.053	.027
	61	.903	.097	.049
	62	.895	.105	.053
21	43	.991	.009	.005
	44	.990	.010	.005
	59	.954	.046	.023
	60	.950	.050	.025
	68	.904	.096	.048
	69	.897	.103	.052

المصدر : عن [Daniel (1978)]

تابع : ملحق (١١)

جدول القيم الحرجة $d(n, \alpha''), d(n, \alpha')$ لاختبار الإشارة

n	d	Confidence coefficient	α''	α'
22	49	.991	.009	.005
	50	.990	.010	.005
	66	.954	.046	.023
	67	.950	.050	.025
	76	.902	.098	.049
	77	.895	.105	.053
23	55	.991	.009	.005
	56	.990	.010	.005
	74	.952	.048	.024
	75	.948	.052	.026
	84	.902	.098	.049
	85	.895	.105	.052
24	62	.990	.010	.005
	63	.989	.011	.005
	82	.951	.049	.025
	83	.947	.053	.026
	92	.905	.095	.048
	93	.899	.101	.051
25	69	.990	.010	.005
	70	.989	.011	.005
	90	.952	.048	.024
	91	.948	.052	.026
	101	.904	.096	.048
	102	.899	.101	.051

المصدر : عن [Daniel (1978)]

جدول القيم الحرجة لاختبار

Mann-Whitney-Wilcoxon

n_1	α	n_2-2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	.005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	.01	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2
	.025	0	0	0	0	0	0	1	1	1	1	2	2	2	2	2	3	3	3	3
	.05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	5	5	5	5
	.10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
3	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	.005	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	3	3	4	4
	.01	0	0	0	0	0	1	1	2	2	2	3	3	3	4	4	5	5	5	6
	.025	0	0	0	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9
	.05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	.10	1	2	2	3	4	5	6	6	7	8	9	10	11	11	12	13	14	15	16
4	.001	0	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4
	.005	0	0	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	0	0	0	1	2	2	3	4	4	5	6	6	7	7	8	9	10	10	11
	.025	0	0	1	2	3	4	5	5	6	7	8	9	10	11	12	12	13	14	15
	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	.10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23
5	.001	0	0	0	0	0	0	1	2	2	3	3	4	4	5	6	6	7	8	8
	.005	0	0	0	1	2	2	3	4	5	6	7	8	8	9	10	11	12	13	14
	.01	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	.025	0	1	2	3	4	6	7	8	9	10	12	13	14	15	16	18	19	20	21
	.05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	.10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
6	.001	0	0	0	0	0	0	2	3	4	5	5	6	7	8	9	10	11	12	13
	.005	0	0	1	2	3	4	5	6	7	8	10	11	12	13	14	16	17	18	19
	.01	0	0	2	3	4	5	7	8	9	10	12	13	14	16	17	19	20	21	23
	.025	0	2	3	4	6	7	9	11	12	14	15	17	18	20	22	23	25	26	28
	.05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	.10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
7	.001	0	0	0	0	1	2	3	4	6	7	8	9	10	11	12	14	15	16	17
	.005	0	0	1	2	4	5	7	8	10	11	13	14	16	17	19	20	22	23	25
	.01	0	1	2	4	5	7	8	10	12	13	15	17	18	20	22	24	25	27	29
	.025	0	2	4	6	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
	.05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	.10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
8	.001	0	0	0	1	2	3	5	6	7	9	10	12	13	15	16	18	19	21	22
	.005	0	0	2	3	5	7	8	10	12	14	16	18	19	21	23	25	27	29	31
	.01	0	1	3	5	7	8	10	12	14	16	18	21	23	25	27	29	31	33	35
	.025	1	3	5	7	9	11	14	16	18	20	23	25	27	30	32	35	37	39	42
	.05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	48
	.10	3	6	8	11	14	17	20	23	25	28	31	34	37	40	43	46	49	52	55

تابع ملحق (١٢)

جدول القيم الحرجة لاختبار

Mann-Whitney-Wilcoxon

n_1	α	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
9	001	0	0	0	2	3	4	6	8	9	11	13	15	16	18	20	22	24	26	27
	005	0	1	2	4	6	8	10	12	14	17	19	21	23	25	28	30	32	34	37
	01	0	2	4	6	8	10	12	15	17	19	22	24	27	29	32	34	37	39	41
	025	1	3	5	8	11	13	16	18	21	24	27	29	32	35	38	40	43	46	49
	05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
10	001	0	0	1	2	4	6	7	9	11	13	15	18	20	22	24	26	28	30	33
	005	0	1	3	5	7	10	12	14	17	19	22	25	27	30	32	35	38	40	43
	01	0	2	4	7	9	12	14	17	20	23	25	28	31	34	37	39	42	45	48
	025	1	4	6	9	12	15	18	21	24	27	30	34	37	40	43	46	49	53	56
	05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
11	001	0	0	1	3	5	7	9	11	13	16	18	21	23	25	28	30	33	35	38
	005	0	1	3	6	8	11	14	17	19	22	25	28	31	34	37	40	43	46	49
	01	0	2	5	8	10	13	16	19	23	26	29	32	35	38	42	45	48	51	54
	025	1	4	7	10	14	17	20	24	27	31	34	38	41	45	48	52	56	59	63
	05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
12	001	0	0	1	3	5	8	10	13	15	18	21	24	26	29	32	35	38	41	43
	005	0	2	4	7	10	13	16	19	22	25	28	32	35	38	42	45	48	52	55
	01	0	3	6	9	12	15	18	22	25	29	32	36	39	43	47	50	54	57	61
	025	2	5	8	12	15	19	23	27	30	34	38	42	46	50	54	58	62	66	70
	05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87
13	001	0	0	2	4	6	9	12	15	18	21	24	27	30	33	36	39	43	46	49
	005	0	2	4	8	11	14	18	21	25	28	32	35	39	43	46	50	54	58	61
	01	1	3	6	10	13	17	21	24	28	32	36	40	44	48	52	56	60	64	68
	025	2	5	9	13	17	21	25	29	34	38	42	46	51	55	60	64	68	73	77
	05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	85
	10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
14	001	0	0	2	4	7	10	13	16	20	23	26	30	33	37	40	44	47	51	55
	005	0	2	5	8	12	16	19	23	27	31	35	39	43	47	51	55	59	64	68
	01	1	3	7	11	14	18	23	27	32	37	41	46	51	56	60	65	70	74	78
	025	2	6	10	14	18	23	27	32	37	42	47	52	57	62	67	72	77	82	87
	05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
15	001	0	0	2	5	8	11	15	18	22	25	29	33	37	41	44	48	52	56	60
	005	0	3	6	9	13	17	21	25	30	34	38	43	47	52	56	61	65	70	74
	01	1	4	8	12	16	20	25	29	34	38	43	48	52	57	62	67	71	76	81
	025	2	6	11	15	20	25	30	35	40	45	50	55	60	65	71	76	81	86	91
	05	4	8	13	19	24	29	34	40	45	51	56	62	67	73	78	84	89	95	101
	10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
16	001	0	0	3	6	9	12	16	20	24	28	32	36	40	44	49	53	57	61	66
	005	0	3	6	10	14	19	23	28	32	37	42	46	51	56	61	66	71	75	80
	01	1	4	8	13	17	22	27	32	37	42	47	52	57	62	67	72	77	83	88
	025	2	7	12	16	22	27	32	38	43	48	54	60	65	71	76	82	87	93	99
	05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120

تابع ملحق (١٢)

جدول القيم المخرجة لاختبار

Mann-Whitney-Wilcoxon

n_1	P	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
17	.001	0	1	3	6	10	14	18	22	26	30	35	39	44	48	53	58	62	67	71
	.005	0	3	7	11	16	20	25	30	35	40	45	50	55	61	66	71	76	82	87
	.01	1	5	9	14	19	24	29	34	39	45	50	56	61	67	72	78	83	89	94
	.025	3	7	12	18	23	29	35	40	46	52	58	64	70	76	82	88	94	100	106
	.05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	.10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
18	.001	0	1	4	7	11	15	19	24	28	33	38	43	47	52	57	62	67	72	77
	.005	0	3	7	12	17	22	27	32	38	43	48	54	59	65	71	76	82	88	93
	.01	1	5	10	15	20	25	31	37	42	48	54	60	66	71	77	83	89	95	101
	.025	3	8	13	19	25	31	37	43	49	56	62	68	75	81	87	94	100	107	113
	.05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	.10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
19	.001	0	1	4	8	12	16	21	26	30	35	41	46	51	56	61	67	72	78	83
	.005	1	4	8	13	18	23	29	34	40	46	52	58	64	70	75	82	88	94	100
	.01	2	5	10	16	21	27	33	39	45	51	57	64	70	76	83	89	95	102	108
	.025	3	8	14	20	26	33	39	46	53	59	66	73	79	86	93	100	107	114	120
	.05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	.10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
20	.001	0	1	4	8	13	17	22	27	33	38	43	49	55	60	66	71	77	83	89
	.005	1	4	9	14	19	25	31	37	43	49	55	61	68	74	80	87	93	100	106
	.01	2	6	11	17	23	29	35	41	48	54	61	68	74	81	88	94	101	108	115
	.025	3	9	15	21	28	35	42	49	56	63	70	77	84	91	99	106	113	120	128
	.05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	.10	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

جدول القيم الحرجة لاختبار Kruskal - Wallis

Sample Sizes					Sample Sizes				
n_1	n_2	n_3	Critical value	α	n_1	n_2	n_3	Critical value	α
2	1	1	2.7000	0.300				4.7000	0.101
2	2	1	3.6000	0.200	4	4	1	6.6667	0.010
2	2	2	4.5714	0.067				6.1667	0.022
			3.7143	0.200				4.9667	0.048
3	1	1	3.2000	0.300				4.8667	0.054
3	2	1	4.2857	0.100				4.1667	0.082
			3.8571	0.133				4.0667	0.102
3	2	2	5.3572	0.029	4	4	2	7.0364	0.006
			4.7143	0.048				6.8727	0.011
			4.5000	0.067				5.4545	0.046
			4.4643	0.105				5.2364	0.052
3	3	1	5.1429	0.043				4.5545	0.098
			4.5714	0.100				4.4455	0.103
			4.000	0.129	4	4	3	7.1439	0.010
3	3	2	6.2500	0.011				7.1364	0.011
			5.3611	0.032				5.5985	0.049
			5.1389	0.061				5.5758	0.051
			4.5556	0.100				4.5455	0.099
			4.2500	0.121				4.4773	0.102
3	3	3	7.2000	0.004	4	4	4	7.6538	0.008
			6.4889	0.011				7.5385	0.011
			5.6889	0.029				5.6923	0.049
			5.6000	0.050				5.6538	0.054
			5.0667	0.086				4.6539	0.097
			4.6222	0.100				4.5001	0.104
4	1	1	3.5714	0.200	5	1	1	3.8571	0.143
4	2	1	4.8214	0.057	5	2	1	5.2500	0.036
			4.5000	0.076				5.0000	0.048
			4.0179	0.114				4.4500	0.071
4	2	2	6.0000	0.014				4.2000	0.095
			5.3333	0.033				4.0500	0.119
			5.1250	0.052	5	2	2	6.5333	0.008
			4.4583	0.100				6.1333	0.013
			4.1667	0.105				5.1600	0.034
4	3	1	5.8333	0.021				5.0400	0.056
			5.2083	0.050				4.3733	0.090
			5.0000	0.057				4.2933	0.122
			4.0556	0.093	5	3	1	6.4000	0.012
			4.8889	0.129				4.9600	0.048
4	3	2	6.4444	0.008				4.8711	0.052
			6.3000	0.011				4.0178	0.095
			5.4444	0.046				3.8400	0.123
			5.4000	0.051	5	3	2	6.9091	0.009
			4.5111	0.098				6.8218	0.010
			4.4444	0.102				5.2509	0.049
4	3	3	6.7455	0.010				5.1055	0.052
			6.7091	0.013				4.6509	0.091
			5.7909	0.046				4.4945	0.101
			5.7273	0.050	5	3	3	7.0788	0.009
			4.7091	0.092				6.9818	0.011

تابع : ملحق (١٣)

جدول القيم الحرجة لاختبار Kruskal - Wallis

Sample Sizes					Sample Sizes				
n_1	n_2	n_3	Critical value	α	n_1	n_2	n_3	Critical value	α
5	3	3	5.6485	0.049	5	5	1	6.8364	0.011
			5.5152	0.051				5.1273	0.046
			4.5333	0.097				4.9091	0.053
			4.4121	0.109				4.1091	0.086
5	4	1	6.9545	0.008				4.0364	0.105
			6.8400	0.011	5	5	2	7.3385	0.010
			4.9855	0.044				7.2692	0.010
			4.8600	0.056				5.3385	0.047
			3.9873	0.098				5.2462	0.051
			3.9600	0.102				4.6231	0.097
5	4	2	7.2045	0.009				4.5077	0.100
			7.1182	0.010	5	5	3	7.5780	0.010
			5.2727	0.049				7.5429	0.010
			5.2682	0.050				5.7055	0.046
			4.5409	0.098				5.6264	0.051
			4.5182	0.101				4.5451	0.100
5	4	3	7.4449	0.010				4.5363	0.102
			7.3949	0.011	5	5	4	7.8229	0.010
			5.6564	0.049				7.7914	0.010
			5.6308	0.050				5.6657	0.049
			4.5487	0.099				5.6429	0.050
			4.5231	0.103				4.5229	0.099
5	4	4	7.7604	0.009				4.5200	0.101
			7.7440	0.011	5	5	5	8.000	0.009
			5.6571	0.049				7.9800	0.010
			5.6176	0.050				5.7800	0.049
			4.6187	0.100				5.6600	0.051
			4.5527	0.102				4.5600	0.100
5	5	1	7.3091	0.009				4.5000	0.102

المصدر : عن [Daniel (1978)]

ملحق (١٤)

جدول القيم المخرجة T_1 السفلي لاختبار الدورات

$B_1 \backslash B_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	3	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

المصدر : عن [Daniel (1978)]

ملحق (١٥)

جدول القيم الحرجة T_2 العليا لاختبار الدورات

$n_2 \backslash n_1$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8					11	12	13	14	14	15	15	16	16	16	16	17	17	17	17
9						13	14	14	15	16	16	17	17	18	18	18	18	18	18
10						13	14	15	16	16	17	17	18	18	18	19	19	19	20
11						13	14	15	16	17	17	18	19	19	19	20	20	20	21
12						13	14	16	16	17	18	19	19	20	20	21	21	21	22
13							15	16	17	18	19	19	20	20	21	21	22	22	23
14							15	16	17	18	19	20	20	21	22	22	23	23	24
15							15	16	18	18	19	20	21	22	22	23	23	24	25
16								17	18	19	20	21	21	22	23	23	24	25	25
17								17	18	19	20	21	22	23	23	24	25	25	26
18								17	18	19	20	21	22	23	24	25	25	26	27
19								17	18	20	21	22	23	23	24	25	26	26	27
20								17	18	20	21	22	23	24	25	25	26	27	28

المصدر : عن [Daniel (1978)]

جدول القيم الحرجة $t_{\alpha, n}^*$ لاختبار سيزمان

n	.001	.005	.010	.025	.050	.100
4	--	--	--	--	.8000	.8000
5	--	--	.9000	.9000	.8000	.7000
6	--	.9429	.8857	.8286	.7714	.6000
7	.9643	.8929	.8571	.7450	.6786	.5357
8	.9286	.8571	.8095	.7143	.6190	.5000
9	.9000	.8167	.7667	.6833	.5833	.4667
10	.8667	.7818	.7333	.6364	.5515	.4424
11	.8364	.7545	.7000	.6091	.5273	.4182
12	.8182	.7273	.6713	.5804	.4965	.3986
13	.7912	.6978	.6429	.5549	.4780	.3791
14	.7670	.6747	.6220	.5341	.4593	.3626
15	.7464	.6536	.6000	.5179	.4429	.3500
16	.7265	.6324	.5824	.5000	.4265	.3382
17	.7083	.6152	.5637	.4853	.4118	.3260
18	.6904	.5975	.5480	.4716	.3994	.3148
19	.6737	.5825	.5333	.4579	.3895	.3070
20	.6586	.5684	.5203	.4451	.3789	.2977
21	.6455	.5545	.5078	.4351	.3688	.2909
22	.6318	.5426	.4963	.4241	.3597	.2829
23	.6186	.5306	.4852	.4150	.3518	.2767
24	.6070	.5200	.4748	.4061	.3435	.2704
25	.5962	.5100	.4654	.3977	.3362	.2646
26	.5856	.5002	.4564	.3894	.3299	.2588
27	.5757	.4915	.4481	.3822	.3236	.2540
28	.5660	.4828	.4401	.3749	.3175	.2490
29	.5567	.4744	.4320	.3685	.3113	.2443
30	.5479	.4665	.4231	.3620	.3059	.2400

المصدر : عن [Daniel (1978)]



Biblioteca Alvalade



0314218